

# Typical behavior of the harmonic measure in critical Galton–Watson trees

Shen LIN \*

*Ecole Normale Supérieure, Paris, France*

*E-mail: shen.lin.math@gmail.com*

## Abstract

We study the typical behavior of the harmonic measure of balls in large critical Galton–Watson trees whose offspring distribution has finite variance. The harmonic measure considered here refers to the hitting distribution of height  $n$  by simple random walk on a critical Galton–Watson tree conditioned to have height greater than  $n$ . We prove that, with high probability, the mass of the harmonic measure carried by a random vertex uniformly chosen from height  $n$  is approximately equal to  $n^{-\lambda}$ , where the constant  $\lambda > 1$  does not depend on the offspring distribution. This universal constant  $\lambda$  is equal to the first moment of the asymptotic distribution of the conductance of size-biased Galton–Watson trees minus 1.

**Keywords:** size-biased Galton–Watson tree, harmonic measure, uniform measure, simple random walk and Brownian motion on trees.

**AMS MSC 2010:** 60J80, 60G50, 60K37.

## 1 Introduction

Let  $\theta$  be a non-degenerate probability measure on  $\mathbb{Z}_+$ , and assume that  $\theta$  has mean one and finite variance  $\sigma^2 > 0$ . Under the probability measure  $\mathbb{P}$ , for every integer  $n \geq 0$ , we let  $\mathsf{T}^{(n)}$  be a Galton–Watson tree with offspring distribution  $\theta$ , conditioned on non-extinction at generation  $n$ . From our assumption that  $\theta$  is critical (i.e. of mean 1),  $\mathsf{T}^{(n)}$  is a.s. a finite tree. We denote by  $\mathsf{T}_n^{(n)}$  the set of all vertices of  $\mathsf{T}^{(n)}$  at generation  $n$ . The classical Yaglom’s theorem in the theory of branching processes states that  $n^{-1} \# \mathsf{T}_n^{(n)}$  converges in distribution to an exponential distribution with parameter  $2/\sigma^2$  (see e.g. Theorem 9.2 in Chapter 1 of [1]).

Conditionally on the tree  $\mathsf{T}^{(n)}$ , we consider a simple random walk on  $\mathsf{T}^{(n)}$  starting from the root. Let  $\Sigma_n$  be the first hitting point of generation  $n$  by the random walk. We call the distribution of  $\Sigma_n$  the harmonic measure  $\mu_n$  at level  $n$ , which is a random probability measure supported on the level set  $\mathsf{T}_n^{(n)}$ . The main result of the present work is the following:

**Theorem 1.** *Let  $\Omega_n$  be a random vertex uniformly chosen from  $\mathsf{T}_n^{(n)}$ . There exists a universal constant  $\lambda \in (1, \infty)$ , which does not depend on the offspring distribution  $\theta$ , such that for every  $\delta > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(n^{-\lambda-\delta} \leq \mu_n(\Omega_n) \leq n^{-\lambda+\delta}\right) = 1. \quad (1)$$

---

\*Supported in part by the grant ANR-14-CE25-0014 (ANR GRAAL)

Loosely speaking, if we look at a typical vertex at level  $n$  of the conditional critical Galton–Watson tree  $\mathsf{T}^{(n)}$ , then as shown by (1), it carries with high probability a mass of order  $n^{-\lambda}$  given by the harmonic measure  $\mu_n$ . Recall that the number of vertices of  $\mathsf{T}^{(n)}$  at generation  $n$  is of order  $n$  according to Yaglom’s theorem. Since  $\lambda > 1$ , our main theorem clearly indicates that the harmonic measure  $\mu_n$  is far from being uniformly spread over  $\mathsf{T}_n^{(n)}$ .

To be more precise, by the definition of  $\Omega_n$ , the convergence (1) can be rewritten as

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{\#\mathsf{T}_n^{(n)}} \sum_{v \in \mathsf{T}_n^{(n)}} \mathbf{1}_{\{\mu_n(v) > n^{-\lambda+\delta} \text{ or } \mu_n(v) < n^{-\lambda-\delta}\}} \right] = 0.$$

Using the theorem of Yaglom, it is easy to see that the preceding convergence is equivalent to

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[ \sum_{v \in \mathsf{T}_n^{(n)}} \mathbf{1}_{\{\mu_n(v) > n^{-\lambda+\delta} \text{ or } \mu_n(v) < n^{-\lambda-\delta}\}} \right] = 0. \quad (2)$$

We take  $\delta \in (0, \lambda - 1)$ , and define  $A_n := \{v \in \mathsf{T}_n^{(n)} : \mu_n(v) > n^{-\lambda+\delta}\}$ . The convergence (2) implies that for every  $\varepsilon > 0$ ,

$$\mathbb{P} \left( \frac{\#A_n}{n} > \varepsilon \right) \xrightarrow[n \rightarrow \infty]{} 0.$$

On the other hand, by the definition of  $A_n$ , we have  $\mu_n(v) \leq n^{-\lambda+\delta}$  for any vertex  $v \in \mathsf{T}_n^{(n)} \setminus A_n$ , and it follows that  $\mu_n(\mathsf{T}_n^{(n)} \setminus A_n) \leq n^{-\lambda+\delta} \#\mathsf{T}_n^{(n)}$ . Using again Yaglom’s theorem, we get that

$$\mathbb{P} \left( \mu_n(\mathsf{T}_n^{(n)} \setminus A_n) > \varepsilon \right) \xrightarrow[n \rightarrow \infty]{} 0.$$

Therefore, it holds with probability tending to 1 as  $n \rightarrow \infty$  that, up to a mass of size  $\varepsilon$ , the harmonic measure  $\mu_n$  is supported on a subset of  $\mathsf{T}_n^{(n)}$  of cardinality smaller than  $\varepsilon n$ . This simple consequence of Theorem 1 has already been observed in a recent paper of Curien and Le Gall [3], where, instead of looking at a vertex typical for the tree, they have considered a vertex typical for the harmonic measure. In Theorem 1 of [3], they have shown the existence of another universal constant  $\beta \in (0, 1)$  independent of the offspring distribution  $\theta$ , such that for every  $\delta > 0$ , we have the convergence in  $\mathbb{P}$ -probability

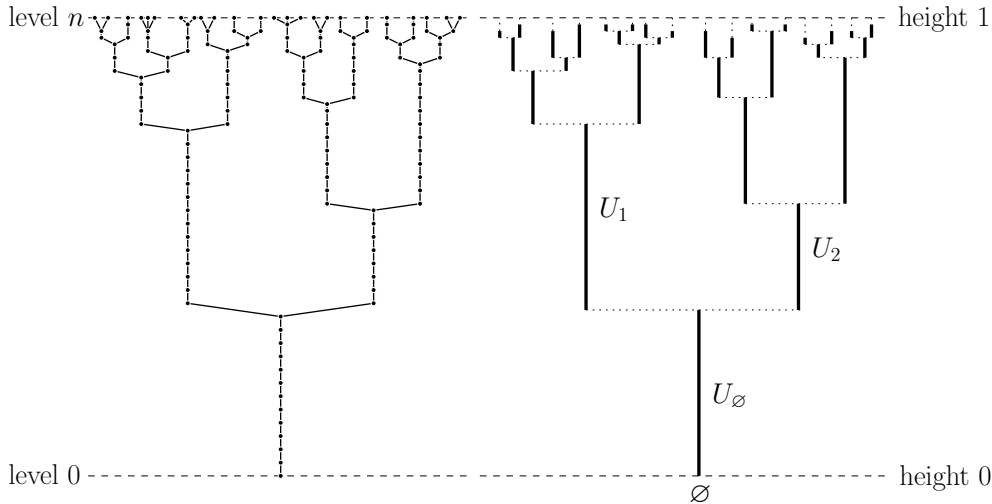
$$\mu_n(\{v \in \mathsf{T}_n^{(n)} : n^{-\beta-\delta} \leq \mu_n(v) \leq n^{-\beta+\delta}\}) \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} 1. \quad (3)$$

In other words, the typical mass given by the harmonic measure  $\mu_n$  to a vertex at level  $n$  drawn with respect to the harmonic measure is with high probability of order  $n^{-\beta}$  with  $\beta < 1$ , and we can thus say that most of the harmonic measure  $\mu_n$  is supported on a subset of size approximately equal to  $n^\beta$ , which is much smaller than the size of  $\mathsf{T}_n^{(n)}$ . One may think of both results, (1) and (3), as the first steps towards a complete analysis of the multifractal nature of the harmonic measure in critical Galton–Watson trees.

As pointed out by Curien and Le Gall in [3], for studying the harmonic measure at level  $n$  on  $\mathsf{T}_n^{(n)}$ , we can consider directly simple random walk on the reduced tree  $\mathsf{T}^{*n}$ , which consists of all vertices of  $\mathsf{T}^{(n)}$  that have at least one descendant at generation  $n$ . Moreover, if we scale the graph distances by  $n^{-1}$ , the rescaled discrete reduced trees  $n^{-1}\mathsf{T}^{*n}$  converge to a random compact rooted  $\mathbb{R}$ -tree  $\Delta$ , whose structure is described as follows. We take a random variable  $U_\emptyset$  uniformly distributed over  $[0, 1]$ , and start with an oriented line segment of length  $U_\emptyset$ ,

whose origin will be the root of  $\Delta$ . To the other end of this initial line segment, we attach the initial points of two new line segments with respective lengths  $U_1$  and  $U_2$ , in such a way that, conditionally given  $U_\emptyset$ , the variables  $U_1$  and  $U_2$  are independent and uniformly distributed over  $[0, 1 - U_\emptyset]$ . To the other end of the first of these 2 line segments, we attach two line segments whose lengths are independent and uniformly distributed over  $[0, 1 - U_\emptyset - U_1]$ , again conditionally on  $U_\emptyset$  and  $U_1$ . We repeat this procedure independently for the second line segment with  $U_1$  replaced by  $U_2$ . We continue this construction by induction, and after an infinite number of steps we obtain a random non-compact rooted  $\mathbb{R}$ -tree  $\Delta_0$  with binary branching (see Figure 1), whose completion with respect to its intrinsic metric  $\mathbf{d}$  is called the continuous reduced tree  $\Delta$ . We assume that  $\Delta$  is also defined under the probability measure  $\mathbb{P}$ . Its boundary  $\partial\Delta$  is the set of all points of  $\Delta$  at height 1, i.e. at distance 1 from the root.

Brownian motion on  $\Delta$  starting from the root can be easily defined up to the first hitting time of  $\partial\Delta$ . Roughly speaking, this process behaves in the same way as a standard linear Brownian motion as long as it remains inside a line segment. It is reflected at the root of the tree  $\Delta$ , and when it arrives at a branching point, it chooses each of the three line segments incident to this point with equal probabilities. The (continuous) harmonic measure  $\mu$  on  $\partial\Delta$  is defined as the (quenched) distribution of the first hitting point of  $\partial\Delta$  by Brownian motion.



**Figure 1:** A large reduced tree  $T^{*n}$  of height  $n$  on the left, and the random tree  $\Delta_0$  on the right

We then define another (non-compact) random rooted  $\mathbb{R}$ -tree  $\Gamma$  with binary branching, such that each point of  $\Gamma$  at height  $y \in [0, \infty)$  corresponds to a point of  $\Delta_0$  at height  $1 - e^{-y} \in [0, 1)$ . The resulting new tree  $\Gamma$  is the Yule tree which describes the genealogy of the classical Yule process, where individuals have independent exponential lifetimes with parameter 1 and each individual has exactly two offspring. By definition, the boundary  $\partial\Gamma$  of  $\Gamma$  is the set of all infinite geodesics in  $\Gamma$  starting from the root (these are called geodesic rays). Due to the binary branching mechanism, both  $\partial\Delta$  and  $\partial\Gamma$  can be canonically identified with  $\{1, 2\}^{\mathbb{N}}$ .

For every  $r > 0$ , we write  $\Gamma_r$  for the level set of  $\Gamma$  at height  $r$ . By a martingale argument, one can define

$$\mathcal{W} := \lim_{r \rightarrow \infty} e^{-r} \# \Gamma_r,$$

and it is well known that  $\mathcal{W}$  follows an exponential distribution of parameter 1. For every  $x \in \Gamma$ ,

we let  $H(x)$  denote the height of  $x$  in  $\Gamma$ , and we write  $\Gamma[x]$  for the tree of descendants of  $x$  in  $\Gamma$ , viewed as an infinite random  $\mathbb{R}$ -tree rooted at  $x$ . For every  $r > 0$ , we write  $\Gamma_r[x]$  for the level set at height  $r$  of the tree  $\Gamma[x]$ . If one thinks of  $\Gamma[x]$  as a subtree of  $\Gamma$ , the set  $\Gamma_r[x]$  consists of all the points of  $\Gamma$  at height  $r + H(x)$  that are descendants of  $x$ . We similarly define

$$\mathcal{W}_x := \lim_{r \rightarrow \infty} e^{-r} \# \Gamma_r[x].$$

It is immediate to see that for every  $r > 0$ ,

$$\sum_{x \in \Gamma_r} e^{-r} \mathcal{W}_x = \mathcal{W}.$$

The uniform measure  $\bar{\omega}$  on  $\partial\Gamma$  is defined as the unique random probability measure on  $\partial\Gamma$  satisfying that, for every  $x \in \Gamma$  and for every geodesic ray  $\mathbf{v} \in \partial\Gamma$  passing through  $x$ ,

$$\bar{\omega}(\mathcal{B}(\mathbf{v}, H(x))) = e^{-H(x)} \frac{\mathcal{W}_x}{\mathcal{W}},$$

where  $\mathcal{B}(\mathbf{v}, H(x))$  stands for the set of all geodesic rays in  $\Gamma$  that coincide with  $\mathbf{v}$  up to height  $H(x)$ . In earlier work,  $\bar{\omega}$  is also named as the branching measure on the boundary of  $\Gamma$ . Recall that  $\partial\Delta$  can be identified with  $\partial\Gamma$  as explained above. We let  $\omega$  be the random probability measure on  $\partial\Delta$  induced by  $\bar{\omega}$ , which will be referred to as the uniform measure on  $\partial\Delta$ .

**Theorem 2.** *With the same constant  $\lambda$  as in Theorem 1, we have  $\mathbb{P}$ -a.s.  $\omega(d\mathbf{v})$ -a.e.*

$$\lim_{r \downarrow 0} \frac{\log \mu(\mathcal{B}_{\mathbf{d}}(\mathbf{v}, r))}{\log r} = \lambda, \quad (4)$$

$$\lim_{r \downarrow 0} \frac{\log \omega(\mathcal{B}_{\mathbf{d}}(\mathbf{v}, r))}{\log r} = 1, \quad (5)$$

where  $\mathcal{B}_{\mathbf{d}}(\mathbf{v}, r)$  stands for the closed ball of radius  $r$  centered at  $\mathbf{v}$  in the metric space  $(\Delta, \mathbf{d})$ .

**Remark.** The Hausdorff measure of  $\partial\Delta$  with respect to  $\mathbf{d}$  is a.s. equal to 1. An exact Hausdorff measure function can be found in Duquesne and Le Gall [4, Theorem 1.3].

**Corollary 3.**  *$\mathbb{P}$ -a.s. the two measures  $\mu$  and  $\omega$  on the boundary of  $\Delta$  are mutually singular.*

In fact, with the same constant  $\beta$  as in (3), it is shown in Theorem 3 of [3] that  $\mathbb{P}$ -a.s.  $\mu(d\mathbf{v})$ -a.e.,

$$\lim_{r \downarrow 0} \frac{\log \mu(\mathcal{B}_{\mathbf{d}}(\mathbf{v}, r))}{\log r} = \beta.$$

If we define

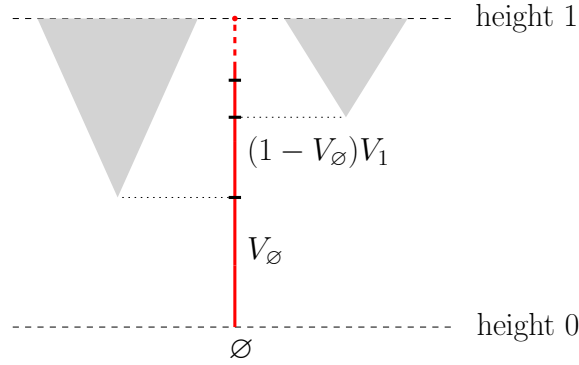
$$B = \{\mathbf{v} \in \partial\Delta : \lim_{r \downarrow 0} \frac{\log \mu(\mathcal{B}_{\mathbf{d}}(\mathbf{v}, r))}{\log r} = \beta\},$$

then  $\mathbb{P}$ -a.s.  $\mu(B) = 1$ . However, since  $\beta < 1 < \lambda$ , we get from (4) that  $\mathbb{P}$ -a.s.  $\omega(B) = 0$ , which finishes the proof of Corollary 3.

Similar results for supercritical infinite Galton–Watson trees can be found in Theorem 3 of Liu and Rouault [7] and in Theorem 6.3 of Lyons, Pemantle and Peres [8], where the uniform measure and the visibility measure (defined as the law of the geodesic ray chosen by *forward* simple random walk) on the boundary of the infinite tree are considered.

In order to get a better understanding of the distinguished geodesic ray in the Yule tree  $\Gamma$  chosen randomly according to the uniform measure  $\bar{\omega}(d\mathbf{v})$ , we follow the ideas of Lyons, Pemantle

and Peres [9] to construct a size-biased version  $\hat{\Gamma}$  of  $\Gamma$ , which is the genealogical tree of the following branching process. Initially, there is one particle having an exponential lifetime with parameter 2, and it reproduces two offspring simultaneously when it dies. We choose one of them uniformly at random and the chosen one will continue as the initial ancestor, while the other offspring will independently evolve as the classical Yule process. The size-biased Yule tree  $\hat{\Gamma}$  thus defined is an infinite random  $\mathbb{R}$ -tree with binary branching. Applying to  $\hat{\Gamma}$  the same transformation that relates  $\Gamma$  and  $\Delta_0$ , we get a bounded (yet non-compact) rooted  $\mathbb{R}$ -tree  $\hat{\Delta}_0$ , which is interpreted as the size-biased version of  $\Delta_0$ . Essentially, every point of  $\hat{\Gamma}$  at height  $y \geq 0$  corresponds to a point of  $\hat{\Delta}_0$  at height  $1 - e^{-y}$ . The completion of  $\hat{\Delta}_0$  with respect to its intrinsic metric is denoted as  $\hat{\Delta}$ , and we call  $\hat{\Delta}$  the size-biased reduced tree. Its boundary  $\partial\hat{\Delta}$  is similarly defined as the set of all points of  $\hat{\Delta}$  at height 1.



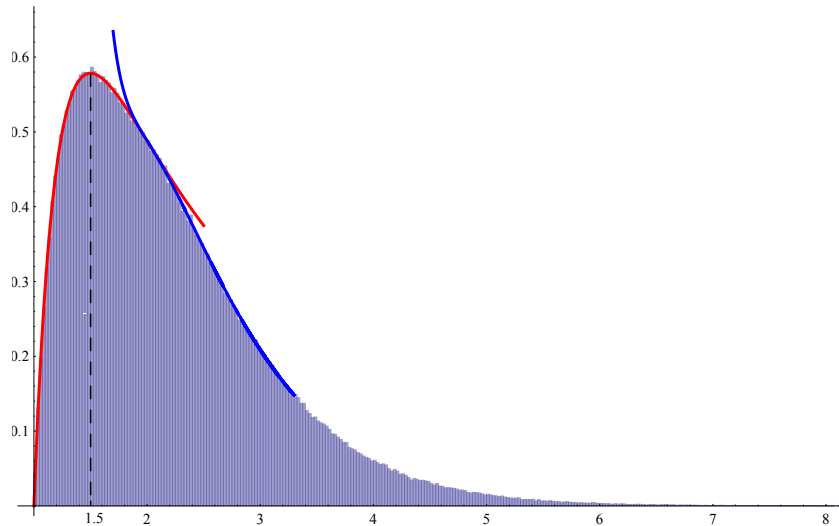
**Figure 2:** Schematic representation of the size-biased reduced tree  $\hat{\Delta}$

Due to the previous description of  $\hat{\Gamma}$ , one can also construct  $\hat{\Delta}$  directly as follows. At first, the root  $\emptyset$  of  $\hat{\Delta}$  has a distinguished descendant line of length 1. Let  $V_\emptyset$  be a random variable taking values in  $[0, 1]$  with density  $2(1 - x)$ , and we graft to the distinguished descendant line at height  $V_\emptyset$  a subtree which is an independent copy of  $\Delta$  scaled by the factor  $(1 - V_\emptyset)$ . In the second step, we take  $V_1$  as an independent copy of  $V_\emptyset$  and graft to the distinguished descendant line at height  $V_\emptyset + (1 - V_\emptyset)V_1$  another independent copy of  $\Delta$  scaled by the factor  $(1 - V_\emptyset)(1 - V_1)$ . Note that for each grafting, we choose the left-hand side or the right-hand side of the distinguished descendant line with equal probabilities. We continue this procedure to graft more subtrees to the distinguished descendant line, with the height of the grafting position increasing to 1. After an infinite number of steps we obtain a realization of  $\hat{\Delta}$ . See Figure 2 for an illustration. We assume as before that  $\hat{\Delta}$  is defined under the probability measure  $\mathbb{P}$ .

The constant  $\lambda$  appearing in Theorems 1 and 2 can be expressed in terms of the (continuous) conductance of  $\hat{\Delta}$ . Informally, if we think of the random trees  $\Delta$  and  $\hat{\Delta}$  as electric networks of resistors with unit resistance per unit length, the effective conductances between the root and the boundary in  $\Delta$  and  $\hat{\Delta}$  are continuous random variables denoted respectively as  $\mathcal{C}$  and  $\hat{\mathcal{C}}$ . From a probabilistic point of view, each of these conductances can be obtained as the mass under the Brownian excursion measure in the corresponding tree for the excursion paths away from the root that hit height 1. It is easy to see that both  $\mathcal{C}$  and  $\hat{\mathcal{C}}$  take values in  $[1, \infty)$ . The law of  $\mathcal{C}$  has been studied at length in [3]. Following the above construction of  $\hat{\Delta}$  and the electric network interpretation, the distribution of  $\hat{\mathcal{C}}$  satisfies the recursive distributional equation

$$\hat{\mathcal{C}} \stackrel{(d)}{=} \left( V + \frac{1 - V}{\hat{\mathcal{C}} + \mathcal{C}} \right)^{-1}, \quad (6)$$

where in the right-hand side  $V, \mathcal{C}$  and  $\widehat{\mathcal{C}}$  are independent, and the distribution of the random variable  $V$  has density  $2(1-x)$  over  $[0, 1]$ . Using this distributional identity (6), we prove that, similarly as the law of  $\mathcal{C}$ , the law  $\widehat{\gamma}$  of  $\widehat{\mathcal{C}}$  has finite moments of all orders, and it has a continuous density  $\widehat{f}$  over  $[1, \infty)$ , which reaches its global maximum at  $3/2$ . The density function  $\widehat{f}$  exhibits a singular behavior analogous to that of the density function of  $\mathcal{C}$  (see [3, Section 2.3]). Although  $\widehat{f}$  is twice continuously differentiable on the interval  $(1, 3)$ , it is shown that  $\widehat{f}$  is not third-order differentiable at the point 2. A similar singular behavior is expected at all integer points  $n \geq 2$ . See Figure 3.



**Figure 3:** A histogram of the distribution  $\widehat{\gamma}$  over  $(1, \infty)$  obtained from the simulations based on (6). The red and the blue curves correspond respectively to the explicit formulae for the density of  $\widehat{\gamma}$  over  $[1, 2]$  and  $[2, 3]$ .

**Proposition 4.** *The distribution  $\widehat{\gamma}$  of the conductance  $\widehat{\mathcal{C}}$  is characterized in the class of all probability measures on  $[1, \infty)$  by the distributional equation (6). The constant  $\lambda$  appearing in Theorems 1 and 2 is given by*

$$\lambda = \mathbb{E}[\widehat{\mathcal{C}}] - 1 \in (1, \infty). \quad (7)$$

Numerical simulations based on (6) and (7) show that  $\lambda \approx 1.21$ .

It is worth pointing out that the main results of [3] have been generalized in [6] to the case where the critical offspring distribution  $\theta$  belongs to the domain of attraction of a stable distribution of index  $\alpha \in (1, 2]$ . We expect that results analogous to those in the present work should hold in the general stable case. This is left to be explored in a future work.

The remainder of this paper is structured as follows. We start by defining formally the continuous random trees  $\Delta$  and  $\Gamma$ . The notation of the random variables involved will be slightly different from the one used in this Introduction. The distribution  $\widehat{\gamma}$  of the conductance  $\widehat{\mathcal{C}}$  is studied in Section 2.4, and the proof of Theorem 2 and of formula (7) is given in Section 2.5. The size-biased continuous random trees  $\widehat{\Delta}$  and  $\widehat{\Gamma}$  are properly defined respectively in Section 2.6 and in Section 2.7. In Section 3, we gather some preliminaries for the discrete setting, where, for example, we prove the convergence of discrete conductances, and introduce a backward version of the discrete size-biased Galton-Watson tree. The last part, Section 4, is devoted to the proof of Theorem 1.

**Acknowledgments.** The author gratefully acknowledges many helpful suggestions of J.-F. Le Gall during the preparation of this work. He is also indebted to N. Curien for several stimulating discussions.

## 2 The continuous setting

### 2.1 The continuous reduced tree $\Delta$

We set

$$\mathcal{V} := \bigcup_{n=0}^{\infty} \{1, 2\}^n,$$

where  $\{1, 2\}^0 = \{\emptyset\}$ . If  $v = (v_1, \dots, v_n) \in \mathcal{V}$ , we set  $|v| = n$  (in particular  $|\emptyset| = 0$ ). If  $v \neq \emptyset$ , we define the parent of  $v$  as  $\bar{v} = (v_1, \dots, v_{n-1})$ , and we then say that  $v$  is a child of  $\bar{v}$ . If both  $u = (u_1, \dots, u_m)$  and  $v = (v_1, \dots, v_n)$  belong to  $\mathcal{V}$ , their concatenation is  $uv := (u_1, \dots, u_m, v_1, \dots, v_n)$ . The notions of a descendant and an ancestor of an element of  $\mathcal{V}$  are defined in the obvious way, with the convention that a vertex  $v \in \mathcal{V}$  is both an ancestor and a descendant of itself. If  $v, w \in \mathcal{V}$ ,  $v \wedge w$  is the unique element of  $\mathcal{V}$  that is an ancestor of both  $v$  and  $w$  and such that  $|v \wedge w|$  is maximal.

We consider a collection  $(U_v)_{v \in \mathcal{V}}$  of independent real random variables uniformly distributed over  $[0, 1]$  under the probability measure  $\mathbb{P}$ . We set  $Y_{\emptyset} = U_{\emptyset}$  and then by induction, for every  $v \in \{1, 2\}^n$  with  $n \geq 1$ , we let

$$Y_v = Y_{\bar{v}} + U_v(1 - Y_{\bar{v}}).$$

Note that a.s.,  $0 \leq Y_v < 1$  for every  $v \in \mathcal{V}$ . Consider the set

$$\Delta_0 := (\{\emptyset\} \times [0, Y_{\emptyset}]) \cup \left( \bigcup_{v \in \mathcal{V} \setminus \{\emptyset\}} \{v\} \times (Y_{\bar{v}}, Y_v] \right).$$

We can define a natural metric  $\mathbf{d}$  on  $\Delta_0$ , so that  $(\Delta_0, \mathbf{d})$  is a (noncompact)  $\mathbb{R}$ -tree and, for every  $x = (v, r) \in \Delta_0$ ,  $\mathbf{d}((\emptyset, 0), x) = r$ . To be specific, let  $x = (v, r) \in \Delta_0$  and  $y = (w, r') \in \Delta_0$ :

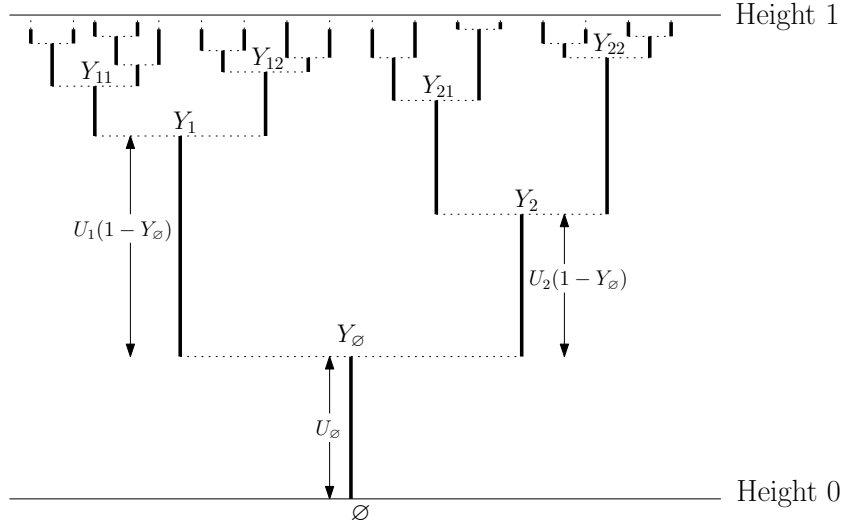
- If  $v$  is a descendant of  $w$  or  $w$  is a descendant of  $v$ , we set  $\mathbf{d}(x, y) = |r - r'|$ .
- Otherwise,  $\mathbf{d}(x, y) = \mathbf{d}((v \wedge w, Y_{v \wedge w}), x) + \mathbf{d}((v \wedge w, Y_{v \wedge w}), y) = (r - Y_{v \wedge w}) + (r' - Y_{v \wedge w})$ .

See Figure 4 for an illustration of the precompact tree  $\Delta_0$ .

We let  $\Delta$  be the completion of  $\Delta_0$  with respect to the metric  $\mathbf{d}$ . Then  $\Delta = \Delta_0 \cup \partial\Delta$ , where the boundary  $\partial\Delta := \{x \in \Delta : \mathbf{d}((\emptyset, 0), x) = 1\}$  is canonically identified with  $\{1, 2\}^{\mathbb{N}}$ . Note that  $(\Delta, \mathbf{d})$  is now a compact  $\mathbb{R}$ -tree. The point  $(\emptyset, 0)$  is called the root of  $\Delta$ . For every  $x \in \Delta$ , we set  $H(x) = \mathbf{d}((\emptyset, 0), x)$  and call it the height of  $x$ . We can define a genealogical order on  $\Delta$  by setting  $x \prec y$  if and only if  $x$  belongs to the geodesic path from the root to  $y$ .

For every  $\varepsilon \in (0, 1)$ , we set  $\Delta_{\varepsilon} := \{x \in \Delta : H(x) \leq 1 - \varepsilon\}$ , which is also a compact  $\mathbb{R}$ -tree for the metric  $\mathbf{d}$ . The leaves of  $\Delta_{\varepsilon}$  are the points of the form  $(v, 1 - \varepsilon)$  for all  $v \in \mathcal{V}$  such that  $Y_{\bar{v}} < 1 - \varepsilon \leq Y_v$ . The branching points of  $\Delta_{\varepsilon}$  are the points of the form  $(v, Y_v)$  for all  $v \in \mathcal{V}$  such that  $Y_v < 1 - \varepsilon$ .

Conditionally on  $\Delta$ , we can take any  $\varepsilon \in (0, 1)$  and define Brownian motion on  $\Delta_{\varepsilon}$  starting from the root. Informally, this process behaves like linear Brownian motion as long as it stays on an “open interval” of the form  $\{v\} \times (Y_{\bar{v}}, Y_v \wedge (1 - \varepsilon))$ , and it is reflected at the root  $(\emptyset, 0)$  and at the leaves of  $\Delta_{\varepsilon}$ . When it arrives at a branching point of the tree, it chooses each of



**Figure 4:** The random tree  $\Delta_0$

three possible line segments ending at this point with equal probabilities. By taking a sequence  $(\varepsilon_n = 2^{-n})_{n \geq 1}$  and then letting  $n$  go to infinity, we can construct under the same probability measure  $P$  a Brownian motion  $(B_t)_{t \geq 0}$  on  $\Delta$  starting from the root up to its first hitting time  $\tau$  of  $\partial\Delta$ . We refer the reader to [3, Section 2.1] for the details of this construction. The harmonic measure  $\mu$  is the distribution of  $B_{\tau-}$  under  $P$ , which is a (random) probability measure on  $\partial\Delta = \{1, 2\}^{\mathbb{N}}$ .

## 2.2 The Yule tree $\Gamma$

To define the Yule tree, consider a collection  $(I_v)_{v \in \mathcal{V}}$  of independent real random variables exponentially distributed with mean 1 under the probability measure  $\mathbb{P}$ . We set  $Z_\emptyset = I_\emptyset$  and then by induction, for every  $v \in \{1, 2\}^n$  with  $n \geq 1$ ,  $Z_v = Z_{\bar{v}} + I_v$ . The Yule tree is the set

$$\Gamma := (\{\emptyset\} \times [0, Z_\emptyset]) \cup \left( \bigcup_{v \in \mathcal{V} \setminus \{\emptyset\}} \{v\} \times (Z_{\bar{v}}, Z_v] \right),$$

which is equipped with the metric  $d$  defined in the same way as  $\mathbf{d}$  in the preceding section. For every  $x = (v, r) \in \Gamma$ ,  $d((\emptyset, 0), x) = r$  and we keep the notation  $H(x) = r$  for the height of the point  $x$ .

Observe that if  $U$  is uniformly distributed over  $[0, 1]$ , the random variable  $-\log(1 - U)$  is exponentially distributed with mean 1. Hence, we may and will suppose that the collection  $(I_v)_{v \in \mathcal{V}}$  is constructed from  $(U_v)_{v \in \mathcal{V}}$  in the previous section via the formula  $I_v = -\log(1 - U_v)$ , for every  $v \in \mathcal{V}$ . Then the mapping  $\Psi$  defined on  $\Delta_0$  by  $\Psi(v, r) = (v, -\log(1 - r))$ , for every  $(v, r) \in \Delta_0$ , is a homeomorphism from  $\Delta_0$  onto  $\Gamma$ .

Using stochastic calculus, we can write, for every  $t \in [0, \tau)$ ,

$$\Psi(B_t) = W \left( \int_0^t (1 - H(B_s))^{-2} ds \right) \quad (8)$$

where  $(W(t))_{t \geq 0}$  is Brownian motion with constant drift  $1/2$  towards infinity on the Yule tree (this process is defined in a similar way as Brownian motion on  $\Delta$ , except that it behaves like



Brownian motion with drift  $1/2$  on every “open interval” of the tree  $\Gamma$ ). Note that  $W$  is also defined under the probability measure  $P$ . From now on, Brownian motion on the Yule tree  $\Gamma$  or on other similar trees will always refer to Brownian motion with drift  $1/2$  towards infinity.

By definition, the boundary  $\partial\Gamma$  is the set of all geodesic rays in  $\Gamma$  starting from the root  $(\emptyset, 0)$ . From the transience of Brownian motion on  $\Gamma$ , there is a.s. a unique geodesic ray denoted by  $W_\infty$  that is visited by  $(W(t), t \geq 0)$  at arbitrarily large times. The distribution of the exit ray  $W_\infty$  under  $P$  yields a probability measure  $\nu$  on the boundary  $\partial\Gamma$ . Thanks to (8), we have in fact  $\nu = \mu$ , provided we identify  $\partial\Delta$  and  $\partial\Gamma$  with  $\{1, 2\}^\mathbb{N}$  and view both  $\mu$  and  $\nu$  as (random) probability measures on  $\{1, 2\}^\mathbb{N}$ .

**Yule-type trees.** We define  $\mathcal{T}$  to be the set of all collections  $(z_v)_{v \in \mathcal{V}}$  of positive real numbers such that the following properties hold:

- (i)  $z_{\bar{v}} < z_v$  for every  $v \in \mathcal{V} \setminus \{\emptyset\}$ ;
- (ii) for every  $\mathbf{v} = (v_1, v_2, \dots) \in \{1, 2\}^\mathbb{N}$ ,  $z_{(v_1, \dots, v_n)} \rightarrow +\infty$  as  $n \rightarrow \infty$ .

Notice that we allow the possibility that  $z_\emptyset = 0$ . If  $(z_v)_{v \in \mathcal{V}} \in \mathcal{T}$ , we consider the associated “tree”

$$T := (\{\emptyset\} \times [0, z_\emptyset]) \cup \left( \bigcup_{v \in \mathcal{V} \setminus \{\emptyset\}} \{v\} \times (z_{\bar{v}}, z_v] \right),$$

which is equipped with the distance  $d$  similarly defined as above. If  $x = (v, r) \in T$ , we still write  $H(x) = r$  for the height of  $x$ . The genealogical (partial) order on  $T$  is defined as previously and will again be denoted by  $\prec$ . The set of all geodesic rays in  $T$  is called the boundary  $\partial T$ , which is naturally identified with  $\{1, 2\}^\mathbb{N}$ . If  $\mathbf{u} = (u_1, u_2, \dots) \in \{1, 2\}^\mathbb{N}$ , and  $x = (v, r) \in T$ , we write  $x \prec \mathbf{u}$  if  $v = (u_1, u_2, \dots, u_k)$  for some integer  $k \geq 0$ .

We will often say that we consider a tree  $T \in \mathcal{T}$ : this means that we are given a collection  $(z_v)_{v \in \mathcal{V}}$  satisfying the above properties, and we consider the associated tree  $T$ . In particular, the tree  $T$  has an order structure given by the lexicographical order on  $\mathcal{V}$ . Elements of  $\mathcal{T}$  will be called Yule-type trees. The Yule tree  $\Gamma$  can be viewed as a random variable taking values in  $\mathcal{T}$ , and we write  $\Theta(dT)$  for its distribution.

Let us fix  $T \in \mathcal{T}$ . If  $r > 0$ , the level set at height  $r$  is  $T_r := \{x \in T : H(x) = r\}$ . If  $x \in T_r$ , we can then consider the subtree  $T[x]$  of descendants of  $x$  in  $T$ . Formally, we view  $T[x]$  as an element of  $\mathcal{T}$ : We write  $v_x$  for the unique element of  $\mathcal{V}$  such that  $x = (v_x, r)$ , and define  $T[x]$  as the Yule-type tree corresponding to the collection  $(z_{v_x v} - r)_{v \in \mathcal{V}}$ .

As we have seen in the Introduction, the limit  $\mathcal{W}(T) = \lim_{r \rightarrow \infty} e^{-r} \#T_r$  exists  $\Theta(dT)$ -a.s., and  $\int \mathcal{W}(T) \Theta(dT) = 1$ . For every  $x \in T$ , we similarly set  $\mathcal{W}(T[x]) = \lim_{r \rightarrow \infty} e^{-r} \#T_r[x]$ . If  $\mathbf{v} \in \partial T$  is a geodesic ray passing through  $x$ , let  $\mathcal{B}(\mathbf{v}, H(x))$  denote the set of geodesic rays in  $T$  that coincide with  $\mathbf{v}$  up to height  $H(x)$ . Then  $\Theta(dT)$ -a.s., the uniform measure  $\bar{\omega}_T$  on  $\partial T$  is defined as the unique probability measure on  $\partial T$  satisfying that

$$\bar{\omega}_T(\mathcal{B}(\mathbf{v}, H(x))) = \frac{e^{-H(x)} \mathcal{W}(T[x])}{\mathcal{W}(T)}, \quad \text{for every } x \in T \text{ and } \mathbf{v} \in \partial T \text{ such that } x \prec \mathbf{v}.$$

On the other hand, for a fixed Yule-type tree  $T \in \mathcal{T}$ , we define the harmonic measure  $\mu_T$  on  $\partial T$  as the distribution of the exit ray chosen by Brownian motion on  $T$  (with drift  $1/2$  towards infinity).

## 2.3 The invariant measure

We write

$$\mathcal{T}^* := \mathcal{T} \times \{1, 2\}^{\mathbb{N}}$$

for the set of all pairs consisting of a tree  $T \in \mathcal{T}$  and a distinguished geodesic ray  $\mathbf{v}$ . Let us define a shift transformation  $S$  on  $\mathcal{T}^*$  by shifting  $(T, \mathbf{v})$  at the first branching point of  $T$ . More precisely, if  $T$  corresponds to the collection  $(z_v)_{v \in \mathcal{V}}$ , we write  $T_{(1)}$  and  $T_{(2)}$  for the two subtrees of  $T$  obtained at the first branching point, which means, for  $i \in \{1, 2\}$ ,  $T_{(i)}$  is the tree corresponding to the collection  $(z_{iv} - z_{\emptyset})_{v \in \mathcal{V}}$ . For any geodesic ray  $\mathbf{v} = (v_1, v_2, \dots)$  in the tree  $T$ , we set  $S(T, \mathbf{v}) := (T_{(v_1)}, \tilde{\mathbf{v}})$ , where  $\tilde{\mathbf{v}} = (v_2, v_3, \dots)$ .

The following proposition is the analogue of Proposition 6.1 in [8] for the Yule tree.

**Proposition 5.** *The probability measure  $\mathcal{W}(T)\Theta(dT)\bar{\omega}_T(d\mathbf{v})$  is invariant under  $S$ .*

*Proof.* Under  $\Theta(dT)$ , if  $T$  corresponds to the collection  $(z_v)_{v \in \mathcal{V}}$ , then  $z_{\emptyset}$  is exponentially distributed with mean 1. Conditionally on  $z_{\emptyset}$ , the branching property of the Yule tree states that  $T_{(1)}$  and  $T_{(2)}$  are i.i.d. of the same law  $\Theta$ .

Let  $F$  be a bounded measurable function on  $\mathcal{T}^*$ . By the definition of the shift  $S$  and the preceding observation, we have

$$\begin{aligned} \int F \circ S(T, \mathbf{v}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{v}) &= \sum_{i=1}^2 \int F(T_{(i)}, \mathbf{u}) (e^{-z_{\emptyset}} \mathcal{W}(T_{(i)})) \Theta(dT) \bar{\omega}_{T_{(i)}}(d\mathbf{u}) \\ &= 2 \int_0^{\infty} e^{-2z} dz \times \int F(T, \mathbf{u}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{u}) \\ &= \int F(T, \mathbf{u}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{u}), \end{aligned}$$

which completes the proof.  $\square$

## 2.4 The continuous conductances

For a fixed Yule-type tree  $T$ , we consider the excursion measure of Brownian motion (with drift  $1/2$ ) on  $T$  away from the root, and define  $\mathcal{C}(T)$  as the mass assigned by this excursion measure to the set of trajectories that never return to the root. Note that  $1 \leq \mathcal{C}(T) < \infty$  for any  $T \in \mathcal{T}$ . For more details on this probabilistic definition of the conductance  $\mathcal{C}(T)$ , we refer the reader to [3, Section 2.3].

To simplify notation, we introduce under the probability measure  $\mathbb{P}$  a pair of random variables  $(\mathcal{W}, \mathcal{C})$  that is distributed as  $(\mathcal{W}(T), \mathcal{C}(T))$  under  $\Theta(dT)$ . In addition, we let  $\hat{\mathcal{C}}$  be a random variable defined under  $\mathbb{P}$  that is distributed as  $\mathcal{C}(T)$  under the probability measure  $\mathcal{W}(T)\Theta(dT)$ .

Let  $T$  be a Yule-type tree corresponding to the collection  $(z_v)_{v \in \mathcal{V}}$ . Recall that  $T_{(1)}$  and  $T_{(2)}$  stand for the two subtrees of  $T$  obtained at the first branching point. From the identity  $\mathcal{W}(T) = e^{-z_{\emptyset}} (\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)}))$  for every  $T \in \mathcal{T}$ , it follows that the distribution of  $\mathcal{W}$  satisfies the distributional equation

$$\mathcal{W} \stackrel{(d)}{=} (1 - U)(\mathcal{W}_1 + \mathcal{W}_2),$$

in which  $\mathcal{W}_1, \mathcal{W}_2$  are two independent copies of  $\mathcal{W}$ , and  $U$  is uniformly distributed over  $[0, 1]$  and independent of  $(\mathcal{W}_1, \mathcal{W}_2)$ . Moreover, the preceding equation holds jointly with a similar distributional identity for the conductance  $\mathcal{C}$  (see equation (2) in [3]). To sum up, we have

$$(\mathcal{W}, \mathcal{C}) \stackrel{(d)}{=} \left( (1 - U)(\mathcal{W}_1 + \mathcal{W}_2), \left( U + \frac{1 - U}{\mathcal{C}_1 + \mathcal{C}_2} \right)^{-1} \right), \quad (9)$$

where  $U$  is as above, while  $(\mathcal{W}_i, \mathcal{C}_i)_{i \in \{1,2\}}$  are two independent copies of  $(\mathcal{W}, \mathcal{C})$ , and are independent of  $U$ .

**Lemma 6.** *The random variable  $\hat{\mathcal{C}}$  satisfies the distributional identity (6).*

*Proof.* By definition, the law of  $\hat{\mathcal{C}}$  is determined by

$$\mathbb{E}[g(\hat{\mathcal{C}})] = \mathbb{E}[\mathcal{W} g(\mathcal{C})] \quad (10)$$

for every nonnegative measurable function  $g$ . Using (9) and symmetry, we have

$$\begin{aligned} \mathbb{E}[g(\hat{\mathcal{C}})] &= \mathbb{E}\left[(1-U)(\mathcal{W}_1 + \mathcal{W}_2) g\left(\left(U + \frac{1-U}{\mathcal{C}_1 + \mathcal{C}_2}\right)^{-1}\right)\right] \\ &= \mathbb{E}\left[2(1-U)\mathcal{W}_1 g\left(\left(U + \frac{1-U}{\mathcal{C}_1 + \mathcal{C}_2}\right)^{-1}\right)\right]. \end{aligned}$$

Recall that the random variable  $V$  in (6) has density function  $2(1-x)$  over  $[0, 1]$ . The statement of the lemma therefore follows by applying (10) in reverse order.  $\square$

The law  $\gamma$  of the conductance  $\mathcal{C}$  has been discussed in great detail in [3, Proposition 6]. We can study the law  $\hat{\gamma}$  of  $\hat{\mathcal{C}}$  by similar arguments. These properties are collected in the next proposition.

For every  $v \in (0, 1)$ ,  $x \geq 0$  and  $c \geq 1$ , we define

$$G(v, x, c) := \left(v + \frac{1-v}{x+c}\right)^{-1}.$$

Let  $\mathcal{M}$  be the set of all probability measures on  $[0, \infty]$  and let  $\hat{\Phi}: \mathcal{M} \rightarrow \mathcal{M}$  map a probability measure  $\sigma$  to

$$\hat{\Phi}(\sigma) = \text{Law}(G(V, X, \mathcal{C})) \quad (11)$$

where  $V$  and  $\mathcal{C}$  are as in (6), while  $X$  is distributed according to  $\sigma$ , and is independent of the pair  $(V, \mathcal{C})$ .

**Proposition 7.** (1) *The distributional equation (6) characterizes the law  $\hat{\gamma}$  of  $\hat{\mathcal{C}}$  in the sense that,  $\hat{\gamma}$  is the unique fixed point of the mapping  $\hat{\Phi}$  on  $\mathcal{M}$ , and for every  $\sigma \in \mathcal{M}$ , the  $k$ -th iterate  $\hat{\Phi}^k(\sigma)$  converges to  $\hat{\gamma}$  weakly as  $k \rightarrow \infty$ .*

(2) *The law  $\hat{\gamma}$  has a continuous density over  $[1, \infty)$ , and all its moments are finite.*

(3) *For any monotone continuously differentiable function  $g: [1, \infty) \rightarrow \mathbb{R}_+$ , we have*

$$\mathbb{E}[\hat{\mathcal{C}}(\hat{\mathcal{C}} - 1)g'(\hat{\mathcal{C}})] + 2\mathbb{E}[g(\hat{\mathcal{C}})] = 2\mathbb{E}[g(\hat{\mathcal{C}} + \mathcal{C})], \quad (12)$$

where  $\hat{\mathcal{C}}$  and  $\mathcal{C}$  are always assumed to be independent under the probability measure  $\mathbb{P}$ .

(4) *We define, for all  $\ell \geq 0$ , the Laplace transforms  $\varphi(\ell) = \mathbb{E}[\exp(-\ell \mathcal{C}/2)]$  and*

$$\hat{\varphi}(\ell) = \mathbb{E}[\exp(-\ell \hat{\mathcal{C}}/2)] = \int_1^\infty e^{-\ell r/2} \hat{\gamma}(dr).$$

*Then  $\hat{\varphi}$  solves the linear differential equation*

$$2\ell \phi''(\ell) + \ell \phi'(\ell) - 2(1 - \varphi(\ell))\phi(\ell) = 0.$$

The proof is very similar to that of the analogous results in [3, Proposition 6]. We therefore skip the details.

**Remark 1.** Using assertion (1) in Proposition 7, one can approximate the law  $\hat{\gamma}$  of  $\hat{\mathcal{C}}$  by iterating the mapping  $\hat{\Phi}$ . An application of the Monte-Carlo method gives  $\mathbb{E}[\hat{\mathcal{C}}] \approx 2.21$ .

**Remark 2.** Following the preceding proposition, we discuss some smoothness properties of the density of  $\hat{\gamma}$ . For every  $t \geq 1$ , we set  $\hat{F}(t) = \mathbb{P}(\hat{\mathcal{C}} \geq t)$ , and we get from (6) that

$$\hat{F}(t) = 2 \left( \frac{t-1}{t} \right)^2 \int_t^\infty dx \frac{x}{(x-1)^3} \mathbb{P}(\hat{\mathcal{C}} + \mathcal{C} \geq x). \quad (13)$$

Since  $\mathbb{P}(\hat{\mathcal{C}} + \mathcal{C} \geq t) = 1$  for every  $t \in [1, 2]$ , we obtain from the last display that

$$\hat{F}(t) = \frac{4t-2}{t^2} A_0 - 2A_0 + 1, \quad \forall t \in [1, 2], \quad (14)$$

where

$$A_0 := 2 - \int_2^\infty dx \frac{x}{(x-1)^3} \mathbb{P}(\hat{\mathcal{C}} + \mathcal{C} \geq x) \in \left( \frac{1}{2}, 2 \right).$$

Let  $\hat{f} = -\hat{F}'$  be the density of the law  $\hat{\gamma}$ . Then it follows from (14) that for all  $t \in [1, 2]$ ,

$$\hat{f}(t) = 4A_0 \times \frac{t-1}{t^3} \quad \text{and} \quad \hat{f}'(t) = 4A_0 \times \frac{3-2t}{t^4}.$$

In particular, we have  $\hat{f}(1) = 0$ ,  $\hat{f}(2) = A_0/2$  and  $\hat{f}'(\frac{3}{2}) = 0$ . Numerical approximations of  $\hat{f}(2)$  show that  $A_0 \approx 0.976$ .

For the density  $f$  of the law of  $\mathcal{C}$ , it is shown in [3, Section 2.3] that there exists a constant  $K_0 \in (1, 2)$  such that  $f(t) = K_0 t^{-2}$  for  $t \in [1, 2]$ . The explicit forms of  $f$  and  $\hat{f}$  over  $[1, 2]$  can be used to calculate the probability  $\mathbb{P}(\hat{\mathcal{C}} + \mathcal{C} \geq t)$  for  $t \in [2, 3]$  by convolution. The values of  $\hat{F}$  over  $[2, 3]$  are thus determined via the ordinary differential equation

$$t(t-1) \hat{F}'(t) - 2 \hat{F}(t) = -2 \mathbb{P}(\hat{\mathcal{C}} + \mathcal{C} \geq t), \quad (15)$$

which is a direct consequence of (13). By solving this differential equation, we are able to get an explicit, yet complicated, expression of  $\hat{F}$  over  $[2, 3]$ , in terms of the two (unknown) parameters  $A_0$  and  $K_0$  (numerical approximations of  $f(1)$  show that  $K_0 \approx 1.477$ ). One can then verify that the density  $\hat{f}$  is continuously differentiable on  $(1, 3)$ . Furthermore, it holds that

$$\hat{f}''(2-) = \hat{f}''(2+) = 0,$$

and that  $\hat{f}$  is twice continuously differentiable on  $(1, 3)$ . However,  $\hat{f}$  is not third-order differentiable at the point 2, as one has

$$\hat{f}'''(2-) = \frac{3A_0}{4}, \quad \text{while} \quad \hat{f}'''(2+) = \frac{3A_0}{4} - 4A_0 K_0.$$

This is similar to the singular behavior of the density  $f$  pointed out in [3, Section 2.3], where it is shown that  $f''(2-) \neq f''(2+)$ . One may conjecture that the density  $\hat{f}$  of  $\hat{\gamma}$  is twice continuously differentiable on the whole interval  $(1, \infty)$ , but not third-order differentiable at all integers  $n \geq 2$ .

We finally remark that  $3/2$  is the global maximum point for the density  $\hat{f}$ . In fact, we have seen that  $\hat{f}$  reaches its maximum at  $3/2$  over the interval  $[1, 2]$ . Meanwhile, it is elementary to verify, by differentiating (15), that the function  $\hat{f}$  is strictly decreasing over  $[3/2, \infty)$ .

## 2.5 Proof of Theorem 2 and of Proposition 4

We have seen in Proposition 5 that the probability measure  $\mathcal{W}(T)\Theta(dT)\bar{\omega}_T(d\mathbf{v})$  on  $\mathcal{T}^*$  is invariant under the shift  $S$ . Recall that  $\mathcal{W}(T)$  follows an exponential distribution of mean 1 under  $\Theta(dT)$ . Taking into account that  $\mathcal{W}(T) > 0$ ,  $\Theta(dT)$ -a.s., we can then verify, in a similar way as in [6, Proposition 2.6], that the shift  $S$  acting on the probability space  $(\mathcal{T}^*, \mathcal{W}(T)\Theta(dT)\bar{\omega}_T(d\mathbf{v}))$  is ergodic. We shall apply Birkhoff's ergodic theorem to the three functionals defined below.

First, let  $H_n(T, \mathbf{v})$  denote the height of the  $n$ -th branching point on the geodesic ray  $\mathbf{v}$ . One immediately verifies that, for every  $n \geq 1$ ,

$$H_n = \sum_{i=0}^{n-1} H_1 \circ S^i,$$

where  $S^i$  stands for the  $i$ -th iterate of the shift  $S$ . It follows thus from the ergodic theorem that  $\mathcal{W}(T)\Theta(dT)\bar{\omega}_T(d\mathbf{v})$ -a.s.,

$$\frac{1}{n}H_n \xrightarrow[n \rightarrow \infty]{} \int H_1(T, \mathbf{v}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{v}). \quad (16)$$

Since the density  $\mathcal{W}(T)$  is a.s. strictly positive, the latter convergence also holds  $\Theta(dT)\bar{\omega}_T(d\mathbf{v})$ -a.s. By the definition of  $H_1$  and then the branching property of the Yule tree,

$$\begin{aligned} \int H_1(T, \mathbf{v}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{v}) &= \int z_{\emptyset} \mathcal{W}(T) \Theta(dT) \\ &= \int z_{\emptyset} e^{-z_{\emptyset}} (\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})) \Theta(dT) \\ &= 2 \left( \int_0^{\infty} z e^{-2z} dz \right) \times \int \mathcal{W}(T) \Theta(dT) = \frac{1}{2}. \end{aligned} \quad (17)$$

Secondly, for a fixed geodesic ray  $\mathbf{v} = (v_1, v_2, \dots) \in \{1, 2\}^{\mathbb{N}}$ , we let  $\mathbf{x}_{n, \mathbf{v}}$  denote the  $n+1$ -st branching point on the geodesic ray  $\mathbf{v}$ , i.e.  $\mathbf{x}_{n, \mathbf{v}} = ((v_1, \dots, v_n), H_{n+1}(T, \mathbf{v}))$ . We set, for every  $n \geq 1$ , the functional  $F_n(T, \mathbf{v}) := \log \bar{\omega}_T(\{\mathbf{u} \in \partial T : \mathbf{x}_{n, \mathbf{v}} \prec \mathbf{u}\})$ . In particular,

$$F_1(T, \mathbf{v}) = \log \frac{\mathcal{W}(T_{(v_1)})}{\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})}.$$

By the definition of  $\bar{\omega}_T$ , one can check that

$$F_n = \sum_{i=0}^{n-1} F_1 \circ S^i.$$

Using the ergodic theorem again, we have  $\Theta(dT)\bar{\omega}_T(d\mathbf{v})$ -a.s.,

$$\frac{1}{n}F_n \xrightarrow[n \rightarrow \infty]{} \int F_1(T, \mathbf{v}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{v}),$$

in which the limit can be calculated as follows:

$$\begin{aligned}
& \int F_1(T, \mathbf{v}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{v}) \\
&= \sum_{i=1}^2 \int e^{-z_{\varnothing}} \mathcal{W}(T_{(i)}) \log \frac{\mathcal{W}(T_{(i)})}{\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})} \Theta(dT) \\
&= \sum_{i=1}^2 \int e^{-z_{\varnothing}} \mathcal{W}(T_{(i)}) \log \mathcal{W}(T_{(i)}) \Theta(dT) - \int \left( \sum_{i=1}^2 e^{-z_{\varnothing}} \mathcal{W}(T_{(i)}) \log (e^{z_{\varnothing}} \mathcal{W}(T)) \right) \Theta(dT) \\
&= 2 \left( \int_0^{\infty} e^{-2z} dz \right) \times \int \mathcal{W}(T) \log \mathcal{W}(T) \Theta(dT) - \int \mathcal{W}(T) \log (e^{z_{\varnothing}} \mathcal{W}(T)) \Theta(dT) \\
&= - \int z_{\varnothing} \mathcal{W}(T) \Theta(dT). \tag{18}
\end{aligned}$$

Note that we used the fact that  $\int \mathcal{W}(T) |\log \mathcal{W}(T)| \Theta(dT) < \infty$  to derive the last equality. In view of (17), we see that  $\Theta(dT) \bar{\omega}_T(d\mathbf{v})$ -a.s.,  $F_n/n$  converges to  $-\frac{1}{2}$  whereas  $H_n/n$  converges to  $\frac{1}{2}$ . By considering the ratio  $F_n/H_n$  and taking  $n \rightarrow \infty$ , we get that  $\Theta(dT)$ -a.s.  $\bar{\omega}_T(d\mathbf{v})$ -a.e.,

$$\lim_{r \rightarrow \infty} \frac{1}{r} \log \bar{\omega}_T(\mathcal{B}(\mathbf{v}, r)) = -1,$$

from which the convergence (5) readily follows.

Thirdly, we turn to the harmonic measure  $\mu_T$  and set, for every  $n \geq 1$ , the functional  $G_n(T, \mathbf{v}) := \log \mu_T(\{\mathbf{u} \in \partial T : \mathbf{x}_{n, \mathbf{v}} \prec \mathbf{u}\})$ . In particular,

$$G_1(T, \mathbf{v}) = \log \frac{\mathcal{C}(T_{(v_1)})}{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})}.$$

The flow property of the harmonic measure  $\mu_T$  (see Lemma 2.3 in [6]) yields that

$$G_n = \sum_{i=0}^{n-1} G_1 \circ S^i.$$

Similarly we have the  $\Theta(dT) \bar{\omega}_T(d\mathbf{v})$ -almost sure convergence

$$\frac{1}{n} G_n \xrightarrow{n \rightarrow \infty} \int G_1(T, \mathbf{v}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{v}), \tag{19}$$

and we calculate the limit

$$\begin{aligned}
& \int G_1(T, \mathbf{v}) \mathcal{W}(T) \Theta(dT) \bar{\omega}_T(d\mathbf{v}) \\
&= \sum_{i=1}^2 \int e^{-z_{\varnothing}} \mathcal{W}(T_{(i)}) \log \frac{\mathcal{C}(T_{(i)})}{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})} \Theta(dT) \\
&= 2 \left( \int_0^{\infty} e^{-2z} dz \right) \times \int \mathcal{W}(T_{(1)}) \log \frac{\mathcal{C}(T_{(1)})}{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})} \Theta(dT) \\
&= \int \mathcal{W}(T_{(1)}) \log \frac{\mathcal{C}(T_{(1)})}{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})} \Theta(dT).
\end{aligned}$$

Putting the convergence (19) together with (16) and (17), we see that  $\Theta(dT)$ -a.s.  $\bar{\omega}_T(d\mathbf{v})$ -a.e.,

$$\lim_{r \rightarrow \infty} \frac{1}{r} \log \mu_T(\mathcal{B}(\mathbf{v}, r)) = 2 \int \mathcal{W}(T_{(1)}) \log \frac{\mathcal{C}(T_{(1)})}{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})} \Theta(dT).$$

Using the branching property of the Yule tree and recalling the notation in Section 2.4, we have therefore  $\mathbb{P}$ -a.s.  $\omega(\mathrm{d}\mathbf{v})$ -a.e. that

$$\lim_{r \downarrow 0} \frac{\log \mu(\mathcal{B}_{\mathbf{d}}(\mathbf{v}, r))}{\log r} = 2 \mathbb{E} \left[ \log \left( \frac{\widehat{\mathcal{C}} + \mathcal{C}}{\widehat{\mathcal{C}}} \right) \right],$$

where  $\widehat{\mathcal{C}}$  and  $\mathcal{C}$  are supposed to be independent under the probability measure  $\mathbb{P}$ . However, by taking  $g(x) = \log(x)$  in (12), we see that

$$\mathbb{E}[\widehat{\mathcal{C}}] - 1 = 2 \mathbb{E} \left[ \log \left( \frac{\widehat{\mathcal{C}} + \mathcal{C}}{\widehat{\mathcal{C}}} \right) \right]. \quad (20)$$

We define  $\lambda := \mathbb{E}[\widehat{\mathcal{C}}] - 1$ . The proof of the convergence (4) is hence completed.

Finally, in view of Proposition 7, it only remains to verify that  $\lambda > 1$ . In fact, we know from the display following (19) that

$$\lambda = 2 \int e^{-z\varnothing} \left( \sum_{i=1}^2 \mathcal{W}(T_{(i)}) \log \frac{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})}{\mathcal{C}(T_{(i)})} \right) \Theta(\mathrm{d}T).$$

By concavity of the logarithm,

$$\sum_{i=1}^2 \frac{\mathcal{W}(T_{(i)})}{\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})} \log \left( \frac{\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})}{\mathcal{W}(T_{(i)})} \cdot \frac{\mathcal{C}(T_{(i)})}{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})} \right) \leq 0,$$

which entails that

$$\sum_{i=1}^2 \mathcal{W}(T_{(i)}) \log \frac{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})}{\mathcal{C}(T_{(i)})} \geq \sum_{i=1}^2 \mathcal{W}(T_{(i)}) \log \frac{\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})}{\mathcal{W}(T_{(i)})}.$$

Notice that the previous inequality is strict if and only if for  $i \in \{1, 2\}$ ,

$$\frac{\mathcal{W}(T_{(i)})}{\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})} \neq \frac{\mathcal{C}(T_{(i)})}{\mathcal{C}(T_{(1)}) + \mathcal{C}(T_{(2)})}.$$

Since the latter property holds with positive probability under  $\Theta(\mathrm{d}T)$ , we have

$$\begin{aligned} \lambda &> 2 \int e^{-z\varnothing} \left( \sum_{i=1}^2 \mathcal{W}(T_{(i)}) \log \frac{\mathcal{W}(T_{(1)}) + \mathcal{W}(T_{(2)})}{\mathcal{W}(T_{(i)})} \right) \Theta(\mathrm{d}T) \\ &= -2 \int F_1(T, \mathbf{v}) \mathcal{W}(T) \Theta(\mathrm{d}T) \bar{\omega}_T(\mathrm{d}\mathbf{v}). \end{aligned}$$

By (18) and (17), the right-hand side of the last display is equal to 1. We have therefore finished the proof of Theorem 2 and of Proposition 4.

## 2.6 The size-biased Yule tree $\widehat{\Gamma}$

Let  $(\widehat{\Gamma}, \widehat{\mathbf{v}})$  denote a random element in  $\mathcal{T}^*$  distributed according to  $\mathcal{W}(T) \Theta(\mathrm{d}T) \bar{\omega}_T(\mathrm{d}\mathbf{v})$ . We give here a direct construction of  $(\widehat{\Gamma}, \widehat{\mathbf{v}}) \in \mathcal{T}^*$  under the probability measure  $\mathbb{P}$ . In the following description, all the random variables involved are supposed to be defined under  $\mathbb{P}$ .

First, we introduce a sequence  $(a_k)_{k \geq 1}$  of i.i.d. random variables uniformly distributed over  $\{1, 2\}$ , and another sequence  $(J_k)_{k \geq 1}$  of i.i.d. real random variables exponentially distributed with mean  $1/2$ . Let  $(\Gamma^{(k)})_{k \geq 1}$  be a collection of independent Yule trees, each of which corresponding respectively to the collection  $(Z_v^{(k)})_{v \in \mathcal{V}}$  with the notation introduced in Section 2.2. We assume that  $(a_k)_{k \geq 1}$ ,  $(J_k)_{k \geq 1}$  and  $(\Gamma^{(k)})_{k \geq 1}$  are independent.

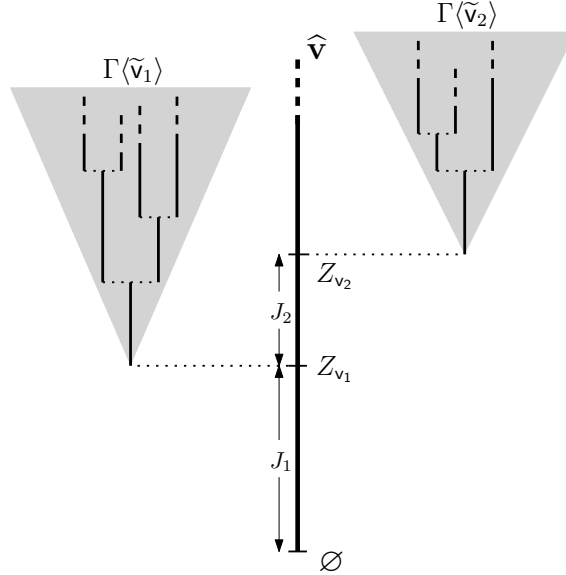
For every integer  $n \geq 1$ , we set  $\mathbf{v}_n = (a_1, a_2, \dots, a_n) \in \{1, 2\}^n$  and  $Z_{\mathbf{v}_n} = \sum_{k=1}^n J_k$ . We write  $\tilde{\mathbf{v}}_n = (a_1, a_2, \dots, a_{n-1}, 3 - a_n) \in \{1, 2\}^n$  for the unique sibling of  $\mathbf{v}_n$  in  $\mathcal{V}$ , and define the subtree  $\Gamma\langle\tilde{\mathbf{v}}_n\rangle$  grafted at  $\tilde{\mathbf{v}}_n$  as

$$\Gamma\langle\tilde{\mathbf{v}}_n\rangle := \left( \{\tilde{\mathbf{v}}_n\} \times (Z_{\mathbf{v}_n}, Z_{\mathbf{v}_n} + Z_{\emptyset}^{(n)}) \right) \cup \left( \bigcup_{v \in \mathcal{V} \setminus \{\emptyset\}} \{\tilde{\mathbf{v}}_n v\} \times (Z_{\mathbf{v}_n} + Z_v^{(n)}, Z_{\mathbf{v}_n} + Z_v^{(n)}) \right).$$

Finally, let  $\hat{\Gamma}$  be the following Yule-type tree

$$\hat{\Gamma} := (\{\emptyset\} \times [0, Z_{\mathbf{v}_1}]) \cup \left( \bigcup_{n \geq 1} \{\mathbf{v}_n\} \times (Z_{\mathbf{v}_n}, Z_{\mathbf{v}_{n+1}}] \right) \cup \left( \bigcup_{n \geq 1} \Gamma\langle\tilde{\mathbf{v}}_n\rangle \right).$$

We will call  $\hat{\Gamma}$  the size-biased Yule tree. See Figure 5 for an illustration.



**Figure 5:** Schematic representation of the size-biased Yule tree  $\hat{\Gamma}$

**Lemma 8.** *The pair  $(\hat{\Gamma}, \hat{\mathbf{v}} = (a_1, a_2, \dots)) \in \mathcal{T} \times \{1, 2\}^{\mathbb{N}}$  constructed above follows the required distribution  $\mathcal{W}(T)\Theta(dT)\bar{\omega}_T(d\mathbf{v})$ .*

The proof of this lemma is based on similar calculations carried out in the previous section. We leave the details to the reader. Note that the analog for discrete-time Galton–Watson trees can be found in Exercise 16.9 in [10].

## 2.7 The size-biased reduced tree $\hat{\Delta}$

Recall the bijection  $\Psi: (v, r) \in \Delta_0 \mapsto (v, -\log(1 - r)) \in \Gamma$  introduced in Section 2.2. We now apply the inverse mapping  $\Psi^{-1}(v, s) = (v, 1 - e^{-s})$  to the size-biased Yule tree  $\hat{\Gamma}$ .



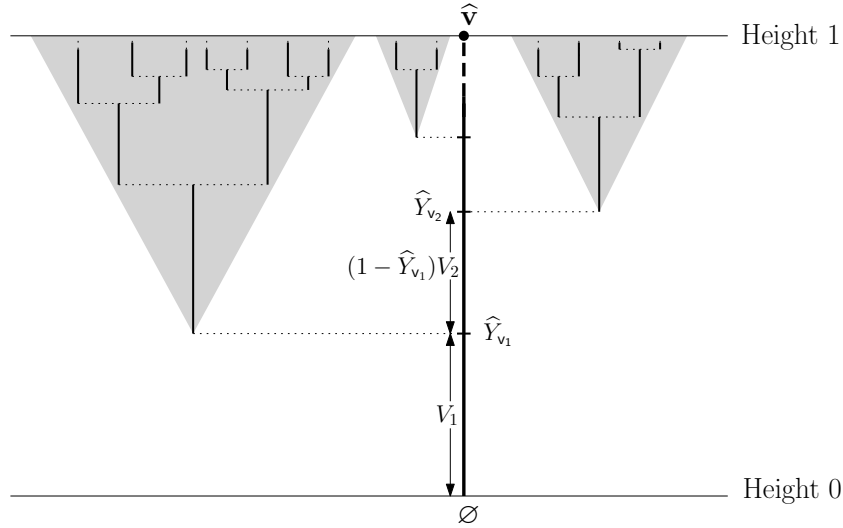
We keep the notation of the preceding section. For every integer  $n \geq 1$ , we set  $V_n = 1 - \exp(-J_n)$ , and then by induction,

$$\hat{Y}_{v_n} = \hat{Y}_{v_{n-1}} + (1 - \hat{Y}_{v_{n-1}})V_n.$$

Notice that  $(V_n)_{n \geq 1}$  are i.i.d. real random variables with density function  $2(1-x)\mathbf{1}_{x \in [0,1]}$ , and that for every  $n \geq 1$ ,  $\hat{Y}_{v_n} = 1 - \exp(-Z_{v_n})$ . Thus,

$$\Psi^{-1}(\hat{\Gamma}) = (\{\emptyset\} \times [0, \hat{Y}_{v_1}]) \cup \left( \bigcup_{n \geq 1} \{v_n\} \times (\hat{Y}_{v_n}, \hat{Y}_{v_{n+1}}] \right) \cup \left( \bigcup_{n \geq 1} \Psi^{-1}(\Gamma(\tilde{v}_n)) \right).$$

We point out that, independently for every  $n \geq 1$ ,  $\Psi^{-1}(\Gamma(\tilde{v}_n))$  is a rescaled copy of the precompact reduced tree  $\Delta_0$  with the scaling factor  $(1 - \hat{Y}_{v_n})$ . From now on we will denote  $\Psi^{-1}(\hat{\Gamma})$  by  $\hat{\Delta}_0$ . See Figure 6 for an illustration of  $\hat{\Delta}_0$ .



**Figure 6:** Schematic representation of the random tree  $\hat{\Delta}_0$

One can define, as for  $\Delta_0$ , the intrinsic metric  $\mathbf{d}$  on  $\hat{\Delta}_0$  such that  $(\hat{\Delta}_0, \mathbf{d})$  is a noncompact  $\mathbb{R}$ -tree, and for every  $x = (v, r) \in \hat{\Delta}_0$ , we have  $\mathbf{d}((\emptyset, 0), x) = r$ . Then we let  $\hat{\Delta}$  be the completion of  $\hat{\Delta}_0$  with respect to  $\mathbf{d}$ , so that  $(\hat{\Delta}, \mathbf{d})$  is a compact  $\mathbb{R}$ -tree. In fact,  $\hat{\Delta} = \hat{\Delta}_0 \cup \partial\hat{\Delta}$ , and the boundary  $\partial\hat{\Delta} := \{x \in \hat{\Delta} : \mathbf{d}((\emptyset, 0), x) = 1\}$  is canonically identified with  $\{1, 2\}^{\mathbb{N}}$ . We will call  $\hat{\Delta}$  the size-biased reduced tree. We keep the same notation  $H(x) = \mathbf{d}((\emptyset, 0), x)$  for the height of  $x \in \hat{\Delta}$ . For every  $\varepsilon \in (0, 1)$ , we set the truncation of  $\hat{\Delta}$  at height  $1 - \varepsilon$

$$\hat{\Delta}_\varepsilon := \{x \in \hat{\Delta} : H(x) \leq 1 - \varepsilon\}.$$

One can think of both  $\Delta$  and  $\hat{\Delta}$  as electric networks of ideal resistors with unit resistance per unit length, and define  $\mathcal{C}(\Delta)$  (resp.  $\mathcal{C}(\hat{\Delta})$ ) be the effective conductance between the root and the set  $\partial\Delta$  (resp.  $\partial\hat{\Delta}$ ) in the corresponding network. As explained in [3, Section 2.3],  $\mathcal{C}(\Delta)$  is identically distributed as the random variable  $\mathcal{C}$  defined in Section 2.4. Analogously,  $\mathcal{C}(\hat{\Delta})$  has the same distribution as  $\hat{\mathcal{C}}$  according to Lemma 8. We thus call  $\hat{\mathcal{C}}$  the continuous conductance of the size-biased reduced tree  $\hat{\Delta}$ .

### 3 The discrete setting

#### 3.1 Notation for discrete trees

We set

$$\mathcal{U} := \bigcup_{n=0}^{\infty} \mathbb{N}^n,$$

where  $\mathbb{N} = \{1, 2, \dots\}$  and  $\mathbb{N}^0 = \{\emptyset\}$  by convention. If  $u = (u_1, \dots, u_n) \in \mathcal{U}$ ,  $|u| = n$  is called the generation (or height) of  $u$ . In particular,  $|\emptyset| = 0$ . A (rooted ordered) tree  $\mathcal{T}$  is a subset of  $\mathcal{U}$  such that the following holds:

- (i)  $\emptyset \in \mathcal{T}$ ;
- (ii) If  $u = (u_1, \dots, u_n) \in \mathcal{T} \setminus \{\emptyset\}$ , then  $\bar{u} := (u_1, \dots, u_{n-1}) \in \mathcal{T}$ ;
- (iii) For every  $u = (u_1, \dots, u_n) \in \mathcal{T}$ , there exists an integer  $k_u(\mathcal{T}) \geq 0$  such that, for every  $j \in \mathbb{N}$ ,  $(u_1, \dots, u_n, j) \in \mathcal{T}$  if and only if  $1 \leq j \leq k_u(\mathcal{T})$ .

The notions of a child and a parent of a vertex of  $\mathcal{T}$  are defined in an obvious way. We write  $\prec$  for the genealogical order on  $\mathcal{T}$ . The quantity  $k_u(\mathcal{T})$  in (iii) is called the number of children of  $u$  in  $\mathcal{T}$ . We always view a tree  $\mathcal{T}$  as a graph whose vertices are the elements of  $\mathcal{T}$  and whose edges are the pairs  $\{\bar{u}, u\}$  for all  $u \in \mathcal{T} \setminus \{\emptyset\}$ .

If  $\mathcal{T}$  is finite, we call it a plane tree. The set of all plane trees is denoted by  $\mathbb{T}_f$ . For an infinite tree  $\mathcal{T}$ , we say it has a single infinite line of descent if there exists a unique sequence of positive integers  $(u_n)_{n \geq 1}$  such that  $(u_1, u_2, \dots, u_n) \in \mathcal{T}$  for all  $n \geq 1$ . We denote by  $\mathbb{T}_\infty$  the set of all infinite trees that have a single infinite line of descent.

The height of a tree  $\mathcal{T}$  is written as  $h(\mathcal{T}) := \sup\{|u| : u \in \mathcal{T}\}$ . The set of all vertices of  $\mathcal{T}$  at generation  $n$  is denoted by  $\mathcal{T}_n := \{u \in \mathcal{T} : |u| = n\}$ . If  $u \in \mathcal{T}$ , the subtree of descendants of  $u$  is  $\tilde{\mathcal{T}}[u] := \{u' \in \mathcal{T} : u \prec u'\}$ . Note that  $\tilde{\mathcal{T}}[u]$  is not a tree under our definition, but we can relabel its vertices to turn it into a tree, by setting  $\mathcal{T}[u] := \{w \in \mathcal{U} : uw \in \mathcal{T}\}$ .

Let  $\mathcal{T}$  be a tree of height larger than  $n$ , and consider a simple random walk  $X = (X_k)_{k \geq 0}$  on  $\mathcal{T}$  starting from the root  $\emptyset$ , which is defined under the probability measure  $P^\mathcal{T}$ . We write  $\tau_n := \inf\{k \geq 0 : |X_k| = n\}$  for the first hitting time of generation  $n$  by  $X$ , and we define the discrete harmonic measure  $\mu_n^\mathcal{T}$  supported on  $\mathcal{T}_n$  as the law of  $X_{\tau_n}$  under  $P^\mathcal{T}$ .

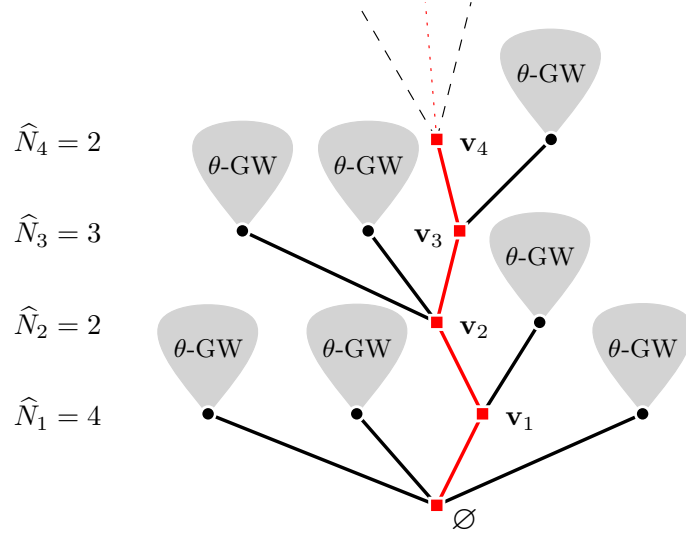
**Critical Galton–Watson trees.** Let  $\theta$  be a non-degenerate probability measure on  $\mathbb{Z}_+$ , and assume that  $\theta$  has mean one and finite variance  $\sigma^2 > 0$ . For every integer  $n \geq 0$ , we let  $\mathsf{T}^{(n)}$  be a Galton–Watson tree with offspring distribution  $\theta$ , conditioned on non-extinction at generation  $n$ , viewed as a random element in  $\mathbb{T}_f$ . In particular,  $\mathsf{T}^{(0)}$  is just a Galton–Watson tree with offspring distribution  $\theta$ . We suppose that the random trees  $\mathsf{T}^{(n)}$  are defined under the probability measure  $\mathbb{P}$ .

Let  $\mathsf{T}^{*n}$  be the reduced tree associated with  $\mathsf{T}^{(n)}$ , which is the random tree composed of all vertices of  $\mathsf{T}^{(n)}$  that have descendants at generation  $n$ . It is always implicitly assumed that we have relabeled the vertices of  $\mathsf{T}^{*n}$ , preserving both the lexicographical order and the genealogical order, so that  $\mathsf{T}^{*n}$  becomes a plane tree in the sense of our preceding definition.

For every  $n \geq 1$  we set  $q_n := \mathbb{P}(\#\mathsf{T}_n^{(0)} > 0)$ . By a standard result (see e.g. Theorem 9.1 of [1, Chapter 1]) on the non-extinction probability up to generation  $n$ , we have

$$q_n \sim \frac{2}{n\sigma^2}, \quad \text{as } n \rightarrow \infty. \quad (21)$$

**Size-biased Galton–Watson tree.** Let  $\hat{N}$  be a random variable distributed according to the size-biased distribution of  $\theta$ , that is, for every  $k \geq 0$ ,  $\mathbb{P}(\hat{N} = k) = k \theta(k)$ . Take a sequence  $(\hat{N}_k)_{k \geq 1}$  of independent copies of  $\hat{N}$  defined under  $\mathbb{P}$ . Now we follow Kesten [5] and Lyons, Pemantle and Peres [9] to construct a size-biased Galton–Watson tree  $\hat{\mathcal{T}}$  defined under  $\mathbb{P}$ . First, the root  $\emptyset$  of  $\hat{\mathcal{T}}$  is given a number  $\hat{N}_1$  of children. Choose one of these children uniformly at random, say  $\mathbf{v}_1$ . It has a number  $\hat{N}_2$  of children, whereas the other children of the root have independently ordinary  $\theta$ -Galton–Watson descendant trees. Again, among the children of  $\mathbf{v}_1$  we choose one uniformly at random, call it  $\mathbf{v}_2$ , and give the others independent  $\theta$ -Galton–Watson descendant trees. Meanwhile the vertex  $\mathbf{v}_2$  has a number  $\hat{N}_3$  of children. Since a.s.  $\hat{N} \geq 1$ , we can repeat this procedure infinitely many times. The resulting random infinite tree  $\hat{\mathcal{T}}$  is called a size-biased Galton–Watson tree (see Figure 7). It is clear by the construction that  $\hat{\mathcal{T}}$  is a random element in  $\mathbb{T}_\infty$  and that its unique infinite line of descent is  $(\mathbf{v}_1, \mathbf{v}_2, \dots)$ , which we will call the spine of  $\hat{\mathcal{T}}$ .



**Figure 7:** Schematic representation of a size-biased Galton–Watson tree  $\hat{\mathcal{T}}$

Let  $[\hat{\mathcal{T}}]^{(n)}$  be the plane tree obtained from  $\hat{\mathcal{T}}$  by keeping only its first  $n$  generations, i.e.,

$$[\hat{\mathcal{T}}]^{(n)} := \{u \in \hat{\mathcal{T}} : |u| \leq n\}.$$

It is shown in [5] and [9] that  $[\hat{\mathcal{T}}]^{(n)}$  is distributed in  $\mathbb{T}_f$  according to the law of  $\mathcal{T}^{(n)}$  biased by  $\#\mathcal{T}_n^{(n)}$ . Moreover, conditionally given the first  $n$  levels of  $\hat{\mathcal{T}}$ , the vertex  $\mathbf{v}_n$  on the spine is uniformly distributed on the  $n$ -th level of  $\hat{\mathcal{T}}$ . Besides, notice that

$$\mathbb{E}[\#\mathcal{T}_n^{(n)}] = \frac{\mathbb{E}[\#\mathcal{T}_n^{(0)}]}{q_n} = \frac{1}{q_n}.$$

All these observations are summarized in the following proposition.

**Proposition 9.** *Let  $F(\mathcal{T}, v)$  be a nonnegative measurable function defined on  $\mathbb{T}_f \times \mathcal{U}$ . Then for every integer  $n \geq 1$ ,*

$$\mathbb{E} \left[ \sum_{v \in \mathcal{T}_n^{(n)}} F(\mathcal{T}^{(n)}, v) \right] = \frac{1}{q_n} \mathbb{E} [F([\hat{\mathcal{T}}]^{(n)}, \mathbf{v}_n)].$$

For every integer  $n \geq 1$ , let  $[\hat{\mathbb{T}}]^n$  be the plane tree obtained from  $\hat{\mathbb{T}}$  by erasing the (infinite) tree of descendants of the vertex  $\mathbf{v}_n$ . By convention, the vertex  $\mathbf{v}_n$  is kept in  $[\hat{\mathbb{T}}]^n$ . Notice that in general  $[\hat{\mathbb{T}}]^n \neq [\hat{\mathbb{T}}]^{(n)}$ , since the height of  $[\hat{\mathbb{T}}]^n$  can be strictly greater than  $n$ .

We let  $[\hat{\mathbb{T}}]^{*n}$  denote the reduced tree associated with the plane tree  $[\hat{\mathbb{T}}]^n$  up to generation  $n$ , which consists of all vertices of  $[\hat{\mathbb{T}}]^n$  that have (at least) one descendant at generation  $n$ . We implicitly assume that the relabelling has been done to turn  $[\hat{\mathbb{T}}]^{*n}$  into a tree. It is elementary to check that  $[\hat{\mathbb{T}}]^{*n}$  is also the reduced tree associated with  $[\hat{\mathbb{T}}]^{(n)}$  up to generation  $n$ .

### 3.2 Convergence of discrete reduced trees

We briefly recall the result in [3] on the convergence of the discrete reduced trees  $\mathbb{T}^{*n}$ . For every real number  $s \in [0, n]$ , we write the truncation of the tree  $\mathbb{T}^{*n}$  at level  $n - \lfloor s \rfloor$  as

$$R_s(\mathbb{T}^{*n}) := \{u \in \mathbb{T}^{*n} : |u| \leq n - \lfloor s \rfloor\}.$$

For every  $\varepsilon \in (0, 1)$ , we have set  $\Delta_\varepsilon = \{x \in \Delta : H(x) \leq 1 - \varepsilon\}$ . We know that, for every fixed  $\varepsilon$ , there is a.s. no branching point of  $\Delta$  at height  $1 - \varepsilon$ . The skeleton of  $\Delta_\varepsilon$  is defined as the following plane tree

$$\text{Sk}(\Delta_\varepsilon) := \{\emptyset\} \cup \{v \in \mathcal{V} \setminus \{\emptyset\} : Y_{\bar{v}} \leq 1 - \varepsilon\} = \{\emptyset\} \cup \{v \in \mathcal{V} \setminus \{\emptyset\} : (\bar{v}, Y_{\bar{v}}) \in \Delta_\varepsilon\}.$$

Consider then the set  $\mathbb{T}_{f, \text{bin}}$  of all plane trees in which every vertex has either 0, 1 or 2 children. For  $\mathcal{T} \in \mathbb{T}_{f, \text{bin}}$  we write  $\mathcal{S}(\mathcal{T})$  for the set of all vertices of  $\mathcal{T}$  having 0 or 2 children. Then there is a unique plane tree  $\langle \mathcal{T} \rangle$  such that one can find a canonical bijection  $u \mapsto w_u$  from  $\langle \mathcal{T} \rangle$  onto  $\mathcal{S}(\mathcal{T})$  that preserves the genealogical order and the lexicographical order of vertices.

The following result is Proposition 16 in [3].

**Proposition 10.** *We can construct the reduced trees  $\mathbb{T}^{*n}$  and the (continuous) tree  $\Delta$  on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  so that the following properties hold for every fixed  $\varepsilon \in (0, 1)$  with  $\mathbb{P}$ -probability one.*

- (i) *For every sufficiently large integer  $n$ , we have  $R_{\varepsilon n}(\mathbb{T}^{*n}) \in \mathbb{T}_{f, \text{bin}}$  and  $\langle R_{\varepsilon n}(\mathbb{T}^{*n}) \rangle = \text{Sk}(\Delta_\varepsilon)$ .*
- (ii) *For every sufficiently large  $n$ , such that the properties stated in (i) hold, and for every  $u \in \text{Sk}(\Delta_\varepsilon)$ , let  $w_u^{n, \varepsilon}$  denote the vertex of  $\mathcal{S}(R_{\varepsilon n}(\mathbb{T}^{*n}))$  corresponding to  $u$  via the canonical bijection from  $\langle R_{\varepsilon n}(\mathbb{T}^{*n}) \rangle$  onto  $\mathcal{S}(R_{\varepsilon n}(\mathbb{T}^{*n}))$ . Then we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} |w_u^{n, \varepsilon}| = Y_u \wedge (1 - \varepsilon).$$

### 3.3 Convergence of discrete conductances

Let  $\mathcal{T} \in \mathbb{T}_f$  be a tree of height larger than  $n$ , and consider the new tree  $\mathcal{T}'$  obtained by adding to the graph  $\mathcal{T}$  an edge between the root  $\emptyset$  and an extra vertex  $\partial$ . We define under the probability measure  $P^{\mathcal{T}'}$  a simple random walk  $X$  on  $\mathcal{T}'$  starting from the root  $\emptyset$ . Let  $\tau_\partial$  be the first hitting time of  $\partial$  by  $X$ , and for every integer  $1 \leq i \leq n$ , let  $\tau_i$  be the first hitting time of generation  $i$  (of the tree  $\mathcal{T}$ ) by  $X$ . We write

$$\mathcal{C}_i(\mathcal{T}) := P^{\mathcal{T}'}(\tau_i < \tau_\partial).$$

This notation is justified by the fact that  $\mathcal{C}_i(\mathcal{T})$  can be interpreted as the effective conductance between  $\partial$  and generation  $i$  of  $\mathcal{T}$  in the graph  $\mathcal{T}'$ , see e.g. [10, Chapter 2].

Recall the notation that  $\mathcal{C}(\Delta)$  stands for the conductance between the root and the set  $\partial\Delta$  in the reduced tree  $\Delta$ . Analogously, for every  $\varepsilon \in (0, 1)$ ,  $\mathcal{C}(\Delta_\varepsilon)$  denotes the conductance between the root and the set  $\{x \in \Delta: H(x) = 1 - \varepsilon\}$  in  $\Delta$ . The following result is stated in [3] without a proof. We provide here the details.

**Proposition 11.** *Suppose that the reduced trees  $\mathsf{T}^{*n}$  and the tree  $\Delta$  are constructed so that the properties stated in Proposition 10 hold. Then*

$$n \mathcal{C}_n(\mathsf{T}^{*n}) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \mathcal{C}(\Delta).$$

*Proof.* By definition, for every  $\varepsilon \in (0, 1)$ ,

$$\mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n}) = P^{(\mathsf{T}^{*n})'}(\tau_{n-\lfloor \varepsilon n \rfloor} < \tau_\partial) \geq P^{(\mathsf{T}^{*n})'}(\tau_n < \tau_\partial) = \mathcal{C}_n(\mathsf{T}^{*n}).$$

Note that there is probability at least  $1 - \frac{\lfloor \varepsilon n \rfloor}{n+1}$  that, after hitting the generation  $n - \lfloor \varepsilon n \rfloor$ , the simple random walk on  $(\mathsf{T}^{*n})'$  will hit the generation  $n$  before moving down to the extra vertex  $\partial$ . Hence it follows from the strong Markov property of simple random walk that

$$0 \leq \mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n}) - \mathcal{C}_n(\mathsf{T}^{*n}) \leq \frac{\lfloor \varepsilon n \rfloor}{n+1} \mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n}).$$

By similar probabilistic arguments, we also have

$$0 \leq \mathcal{C}(\Delta_\varepsilon) - \mathcal{C}(\Delta) \leq \varepsilon \mathcal{C}(\Delta_\varepsilon),$$

which entails particularly that  $\mathcal{C}(\Delta_\varepsilon) \leq 2\mathcal{C}(\Delta)$  if  $\varepsilon < 1/2$ .

Let  $n$  be sufficiently large so that assertions (i) and (ii) of Proposition 10 hold with  $\varepsilon \in (0, \frac{1}{2})$ . By calculating the conductances using the series law and parallel law, we see that a.s.

$$\lim_{n \rightarrow \infty} |n \mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n}) - \mathcal{C}(\Delta_\varepsilon)| = 0.$$

Then, it follows from

$$\begin{aligned} |n \mathcal{C}_n(\mathsf{T}^{*n}) - \mathcal{C}(\Delta)| &\leq |n \mathcal{C}_n(\mathsf{T}^{*n}) - n \mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n})| + |n \mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n}) - \mathcal{C}(\Delta_\varepsilon)| + |\mathcal{C}(\Delta_\varepsilon) - \mathcal{C}(\Delta)| \\ &\leq \lfloor \varepsilon n \rfloor \mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n}) + |n \mathcal{C}_{n-\lfloor \varepsilon n \rfloor}(\mathsf{T}^{*n}) - \mathcal{C}(\Delta_\varepsilon)| + \varepsilon \mathcal{C}(\Delta_\varepsilon) \end{aligned}$$

that

$$\limsup_{n \rightarrow \infty} |n \mathcal{C}_n(\mathsf{T}^{*n}) - \mathcal{C}(\Delta)| \leq 2\varepsilon \mathcal{C}(\Delta_\varepsilon) \leq 4\varepsilon \mathcal{C}(\Delta).$$

Letting  $\varepsilon \rightarrow 0$ , we conclude that  $|n \mathcal{C}_n(\mathsf{T}^{*n}) - \mathcal{C}(\Delta)| \rightarrow 0$  as  $n \rightarrow \infty$ .  $\square$

Let us write  $\mathcal{C}(\widehat{\Delta}_\varepsilon)$  for the conductance between the root and the set  $\{x \in \widehat{\Delta}: H(x) = 1 - \varepsilon\}$  in  $\widehat{\Delta}$ . By the same reasoning as in the previous proof, for every  $\varepsilon \in (0, 1/2)$ , we have

$$0 \leq \mathcal{C}(\widehat{\Delta}_\varepsilon) - \mathcal{C}(\widehat{\Delta}) \leq 2\varepsilon \mathcal{C}(\widehat{\Delta}). \quad (22)$$

**Remark.** We can also show the convergence of the reduced trees  $[\widehat{\mathsf{T}}]^{*n}$  to the (continuous) size-biased reduced tree  $\widehat{\Delta}$ , in a sense similar to Proposition 10, which implies a size-biased analog of Proposition 11. In particular, it holds the convergence in law

$$n \mathcal{C}_n([\widehat{\mathsf{T}}]^n) \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{C}(\widehat{\Delta}).$$

Since these results are not needed in Section 4 for proving Theorem 1, we omit the proofs.

For future reference, we state the following result, which is Lemma 22 in [3].

**Lemma 12.** *There exists a constant  $K \geq 1$  such that, for every integer  $n \geq 1$ ,*

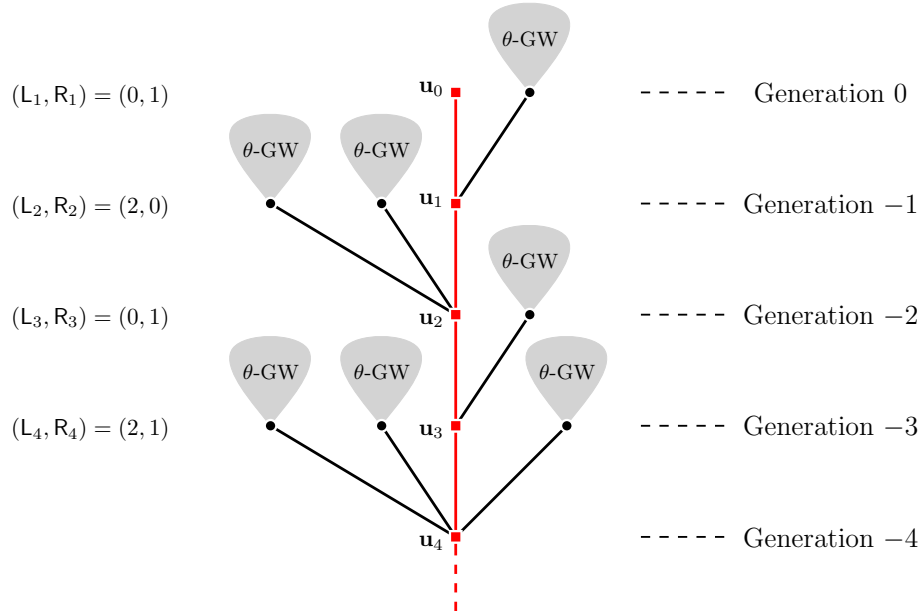
$$\mathbb{E}[(n \mathcal{C}_n(\mathsf{T}^{*n}))^2] \leq K.$$

### 3.4 Backward size-biased Galton–Watson tree

We introduce in this section a new infinite random tree  $\check{\mathcal{T}}$ , which originates from the inflated Galton–Watson tree constructed by Peres and Zeitouni in [11]. It will be clear from the following descriptions that  $\check{\mathcal{T}}$  is a rear-view variant of the size-biased Galton–Watson tree  $\hat{\mathcal{T}}$ .

First, the random tree  $\check{\mathcal{T}}$  has a unique infinite ray of vertices  $(\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \dots)$ , which will be referred to as its spine. We declare that, for every  $n \geq 0$ , the vertex  $\mathbf{u}_n$  is at generation  $-n$ . This gives a genealogical order on the spine of  $\check{\mathcal{T}}$ :  $\mathbf{u}_1$  is viewed as the parent of  $\mathbf{u}_0$ ,  $\mathbf{u}_2$  is viewed as the parent of  $\mathbf{u}_1$ , and so on.

Next, let us describe the finite subtrees in  $\check{\mathcal{T}}$  branching off every node of the spine. To this end, we recall that  $\hat{N}$  follows the size-biased distribution of  $\theta$ , and we denote by  $L$  a random variable which, conditionally on  $\hat{N}$ , is uniformly distributed on the set  $\{0, 1, \dots, \hat{N} - 1\}$ . Let  $(L_n, \hat{N}_n)_{n \geq 1}$  be a sequence of i.i.d. copies of  $(L, \hat{N})$ , and set  $R_n = \hat{N}_n - L_n - 1$  for every  $n \geq 1$ . To every vertex  $\mathbf{u}_n$  we give a number  $L_n$  of children to the left of the spine and a number  $R_n$  of children to the right of the spine. Each of these children (there are  $\hat{N}_n - 1$  in total) will independently have an ordinary  $\theta$ -Galton–Watson descendant tree (later we will say that these Galton–Watson trees are grafted at  $\mathbf{u}_n$ ), and we also assume the independence of these Galton–Watson trees among all  $n \geq 1$ . This finishes the construction of  $\check{\mathcal{T}}$ . See Figure 8 for an illustration. We remark that  $\check{\mathcal{T}}$  is not a tree in the sense of Section 3.1. However, due to its obvious tree structure, we will call  $\check{\mathcal{T}}$  the backward size-biased Galton–Watson tree.



**Figure 8:** Schematic representation of the backward size-biased Galton–Watson tree  $\check{\mathcal{T}}$

The genealogical (partial) order on  $\check{\mathcal{T}}$  is defined in the following way. We simply keep the genealogical orders inherited from the grafted Galton–Watson trees and combine them with the genealogical order on the spine. For instance,  $\mathbf{u}_2$  is an ancestor of any vertex in the subtrees grafted at  $\mathbf{u}_1$ . We can also define in a consistent manner the notion of generation for every vertex in  $\check{\mathcal{T}}$ . In fact, for any vertex  $v$  not on the spine, there is a unique vertex  $\mathbf{u}_m$  on the spine such that  $v$  belongs to a finite subtree grafted at  $\mathbf{u}_m$ , then we say that the generation of  $v$  in  $\check{\mathcal{T}}$

is equal to  $-m + 1$  plus the initial generation of  $v$  inside the corresponding grafted plane tree.

The vertex  $\mathbf{u}_0$  in  $\check{\mathbb{T}}$  corresponds to the root in the inflated Galton–Watson tree (abbreviated as *IGW*) described in Peres and Zeitouni [11]. It is worth pointing out that, while an independent number of offspring of distribution  $\theta$  is assigned the root of *IGW*, the vertex  $\mathbf{u}_0$  in  $\check{\mathbb{T}}$  has no descendants. Besides, our notion of generation on  $\check{\mathbb{T}}$  can also be called as the horocycle distance from  $\mathbf{u}_0$ , if we follow the terminology in [11].

For every  $n \geq 1$ , let  $[\check{\mathbb{T}}]^n$  be the plane tree obtained from  $\check{\mathbb{T}}$  by only keeping the finite tree above the vertex  $\mathbf{u}_n$ . We take  $\mathbf{u}_n$  as the root of  $[\check{\mathbb{T}}]^n$ , and the lexicographical order on the set of vertices of  $[\check{\mathbb{T}}]^n$  corresponds to the order of visit when one “moves around” the finite tree  $[\check{\mathbb{T}}]^n$  in clockwise order, starting from the root  $\mathbf{u}_n$ . A key observation is that, viewed as a random plane tree,  $[\check{\mathbb{T}}]^n$  has the same distribution as the random tree  $[\hat{\mathbb{T}}]^n$  defined in Section 3.1. Moreover, the root  $\mathbf{u}_n$  of  $[\check{\mathbb{T}}]^n$  corresponds to the root  $\emptyset$  of  $[\hat{\mathbb{T}}]^n$ , and the vertex  $\mathbf{u}_0$  in  $[\check{\mathbb{T}}]^n$  corresponds to the vertex  $\mathbf{v}_n$  in  $[\hat{\mathbb{T}}]^n$ .

## 4 Proof of Theorem 1

Before we start the proof of Theorem 1, let us emphasize that Theorem 1 is not a straightforward consequence of Theorem 2. In fact, we will not directly use the results (4) and (5) obtained in the continuous setting.

Recall that  $\lambda = \mathbb{E}[\hat{\mathcal{C}}] - 1$  is the constant greater than 1 that appears in the convergence (4). Let  $\delta > 0$ . By applying Proposition 9 to the indicator function

$$F(\mathcal{T}, v) = \mathbf{1}\{n^{-\lambda-\delta} \leq P^{\mathcal{T}}(X_{\tau_n} = v) \leq n^{-\lambda+\delta}\}^c, \quad \text{for } \mathcal{T} \in \mathbb{T}_f \text{ and } v \in \mathcal{T},$$

we see that

$$\mathbb{E} \left[ \sum_{v \in \mathbb{T}_n^{(n)}} \mathbf{1}\{n^{-\lambda-\delta} \leq P^{\mathbb{T}^{(n)}}(X_{\tau_n} = v) \leq n^{-\lambda+\delta}\}^c \right] = \frac{1}{q_n} \mathbb{P} \left( \{n^{-\lambda-\delta} \leq P^{[\hat{\mathbb{T}}]^{(n)}}(X_{\tau_n} = \mathbf{v}_n) \leq n^{-\lambda+\delta}\}^c \right).$$

Notice that in the left-hand side of the last display,  $P^{\mathbb{T}^{(n)}}(X_{\tau_n} = v)$  is by definition the harmonic measure  $\mu_n(v)$  at vertex  $v$ . By virtue of (2) and (21), the proof of convergence (1) is thus reduced to showing that for every  $\delta > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( n^{-\lambda-\delta} \leq P^{[\hat{\mathbb{T}}]^{(n)}}(X_{\tau_n} = \mathbf{v}_n) \leq n^{-\lambda+\delta} \right) = 1. \quad (23)$$

Since the hitting distribution of generation  $n$  is the same for simple random walk on  $[\hat{\mathbb{T}}]^{(n)}$  and on its reduced tree  $[\hat{\mathbb{T}}]^{*n}$ , we have the equality

$$P^{[\hat{\mathbb{T}}]^{(n)}}(X_{\tau_n} = \mathbf{v}_n) = P^{[\hat{\mathbb{T}}]^{*n}}(X_{\tau_n} = \mathbf{v}_n) = P^{[\hat{\mathbb{T}}]^n}(X_{\tau_n} = \mathbf{v}_n)$$

under the probability measure  $\mathbb{P}$ . Furthermore, according to the final remark in Section 3.4, the (random) probability  $P^{[\hat{\mathbb{T}}]^n}(X_{\tau_n} = \mathbf{v}_n)$  is distributed under  $\mathbb{P}$  as

$$P^{[\check{\mathbb{T}}]^n}(X_{\tau_n} = \mathbf{u}_0),$$

by which we mean the probability that a simple random walk on  $[\check{\mathbb{T}}]^n$  starting from the root  $\mathbf{u}_n$  hits level  $n$  (of  $[\check{\mathbb{T}}]^n$ ) for the first time at  $\mathbf{u}_0$ . So the convergence (23) is equivalent to

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( n^{-\lambda-\delta} \leq P^{[\check{\mathbb{T}}]^n}(X_{\tau_n} = \mathbf{u}_0) \leq n^{-\lambda+\delta} \right) = 1. \quad (24)$$

In order to show the latter convergence, we denote by  $-M_1, -M_2, \dots$  the generations of the vertices on the spine of  $\check{\mathbf{T}}$  where there is (at least) one grafted plane tree that reaches generation 0, i.e. has a descendant of generation 0. This sequence of negative integers  $(-M_k)_{k \geq 1}$  is listed in the strict decreasing order, and we set by convention  $M_0 = 0$ . For every  $k \geq 1$ , we also set  $L_k := M_k - M_{k-1} \geq 1$ .

For every  $n \geq 1$ , let  $k_n := k_n(\check{\mathbf{T}})$  be the index such that  $M_{k_n} \leq n < M_{k_n+1}$ .

**Lemma 13.** *We have  $\mathbb{P}$ -a.s.*

$$\lim_{n \rightarrow \infty} \frac{k_n}{2 \log n} = 1.$$

*Proof.* Recall that for every  $j \geq 1$ , there are  $\hat{N}_j - 1$  independent Galton–Watson trees grafted at  $\mathbf{u}_j$  in  $\check{\mathbf{T}}$ . Consider the event that at least one of those plane trees grafted at  $\mathbf{u}_j$  reaches generation 0, and let  $\epsilon_j$  be the corresponding indicator function. Then,

$$\mathbb{P}(\epsilon_j = 0) = \mathbb{E}[(1 - q_{j-1})^{\hat{N}_j - 1}] = \mathbb{E}[(1 - q_{j-1})^{\hat{N} - 1}].$$

Let  $g_\theta$  be the generating function of  $\theta$ , i.e.

$$g_\theta(r) := \sum_{k \geq 0} \theta(k) r^k, \quad 0 \leq r \leq 1.$$

Since  $\theta$  has a finite variance  $\sigma^2$ ,

$$g_\theta(1 - s) = 1 - s + \frac{\sigma^2}{2} s^2 + o(s^2) \quad \text{as } s \rightarrow 0.$$

As the mean of  $\hat{N} - 1$  is  $\sigma^2$ , we have  $\mathbb{E}[(1 - s)^{\hat{N} - 1}] = 1 - \sigma^2 s + o(s)$  as  $s \rightarrow 0$ , which, together with (21), yields that

$$\mathbb{P}(\epsilon_j = 0) = 1 - \frac{2}{j} + o(j^{-1}) \quad \text{as } j \rightarrow \infty. \quad (25)$$

Notice that by definition,  $k_n = \epsilon_1 + \epsilon_2 + \dots + \epsilon_n$ . Hence,

$$\mathbb{E}[k_n] = \sum_{j=1}^n (1 - \mathbb{P}(\epsilon_j = 0)) \sim 2 \log n \quad \text{as } n \rightarrow \infty.$$

Since  $\epsilon_1, \dots, \epsilon_n$  are independent, we also have  $\text{var}(k_n) = O(\log n)$ , and the  $L^2$ -convergence of  $k_n/(2 \log n)$  follows immediately. The a.s. convergence is then obtained by standard monotonicity and Borel–Cantelli arguments.  $\square$

We introduce some additional notation before stating the next proposition. For every  $j \geq 0$ , we write  $P_j^{\check{\mathbf{T}}}$  for the (quenched) probability measure under which we consider a simple random walk  $X = (X_k)_{k \geq 0}$  on  $\check{\mathbf{T}}$  starting from the vertex  $\mathbf{u}_j$ . Under  $P_j^{\check{\mathbf{T}}}$ , we denote by  $S_0$  the hitting time of generation 0 by the simple random walk  $X$ , and for every  $i \geq 0$ ,  $\Pi_i := \inf\{k \geq 0 : X_k = \mathbf{u}_i\}$  denotes the hitting time of vertex  $\mathbf{u}_i$ .



**Proposition 14.** *For every  $\delta > 0$ , there exists an integer  $n_0 \in \mathbb{N}$  such that for every  $n \geq n_0$ , we have*

$$\mathbb{P}\left(P^{[\check{\Gamma}]^n}(X_{\tau_n} = \mathbf{u}_0) \geq n^{-\lambda+\delta}\right) \leq 8 \mathbb{P}\left(P_{M_{k_n}}^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}}) \geq n^{-\lambda+\delta}/2\right),$$

and

$$\mathbb{P}\left(P^{[\check{\Gamma}]^n}(X_{\tau_n} = \mathbf{u}_0) \leq n^{-\lambda-\delta}\right) \leq 8 \mathbb{P}\left(P_{M_{k_n}}^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}}) \leq n^{-\lambda-\delta}\right).$$

*Proof.* We keep the notation used in the proof of Lemma 13. Observe that

$$\mathbb{P}(M_{k_n+1} - n > n) = \mathbb{P}(\epsilon_{n+j} = 0 \text{ for all } 1 \leq j \leq n) = \prod_{j=1}^n \mathbb{P}(\epsilon_{n+j} = 0)$$

converges to  $1/4$  as  $n \rightarrow \infty$  by (25). One can thus find an integer  $n_0$  such that for every  $n \geq n_0$ ,

$$\mathbb{P}(M_{k_n+1} > 2n) \geq \frac{1}{8}. \quad (26)$$

Since  $M_{k_n+1} - n$  is independent of the finite tree above the vertex  $\mathbf{u}_n$  in  $\check{\Gamma}$ ,

$$\mathbb{P}(M_{k_n+1} > 2n, P^{[\check{\Gamma}]^n}(X_{\tau_n} = \mathbf{u}_0) \geq n^{-\lambda+\delta}) = \mathbb{P}(M_{k_n+1} > 2n) \times \mathbb{P}(P^{[\check{\Gamma}]^n}(X_{\tau_n} = \mathbf{u}_0) \geq n^{-\lambda+\delta}). \quad (27)$$

On the other hand, it is crucial to notice that under the probability measure  $\mathbb{P}$ , the probability  $P^{[\check{\Gamma}]^n}(X_{\tau_n} = \mathbf{u}_0)$  has the same distribution as the conditional probability

$$P_n^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0 | S_0 < \Pi_{M_{k_n+1}}),$$

which can be calculated as

$$\begin{aligned} P_n^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0 | S_0 < \Pi_{M_{k_n+1}}) &= \frac{P_n^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}})}{P_n^{\check{\Gamma}}(S_0 < \Pi_{M_{k_n+1}})} \\ &= \frac{P_n^{\check{\Gamma}}(\Pi_{M_{k_n}} < \Pi_{M_{k_n+1}}) \times P_{M_{k_n}}^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}})}{P_n^{\check{\Gamma}}(S_0 < \Pi_{M_{k_n+1}})} \end{aligned}$$

by the strong Markov property of the random walk. Besides, simple considerations show that  $P_n^{\check{\Gamma}}(S_0 < \Pi_{M_{k_n+1}}) \geq 1/2$  on the event  $\{M_{k_n+1} > 2n\}$ . Hence, we have

$$\begin{aligned} \mathbb{P}(M_{k_n+1} > 2n, P^{[\check{\Gamma}]^n}(X_{\tau_n} = \mathbf{u}_0) \geq n^{-\lambda+\delta}) &= \mathbb{P}(M_{k_n+1} > 2n, P_n^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0 | S_0 < \Pi_{M_{k_n+1}}) \geq n^{-\lambda+\delta}) \\ &\leq \mathbb{P}(M_{k_n+1} > 2n, P_{M_{k_n}}^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}}) \geq n^{-\lambda+\delta}/2) \\ &\leq \mathbb{P}(P_{M_{k_n}}^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}}) \geq n^{-\lambda+\delta}/2). \end{aligned}$$

The last display, together with (26) and (27), yields the first inequality in the statement of the proposition.

We can argue in a similar manner for the second inequality stated in the proposition. Its proof is even simpler because it suffices to use the bound

$$P_{M_{k_n}}^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}}) \leq P_n^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0 | S_0 < \Pi_{M_{k_n+1}}),$$

instead of the estimate  $P_n^{\check{\Gamma}}(S_0 < \Pi_{M_{k_n+1}}) \geq 1/2$  used above.  $\square$

According to (24) and the preceding result, we can therefore derive Theorem 1 from the following proposition.

**Proposition 15.** *For every  $\delta > 0$ , it holds that*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(n^{-\lambda-\delta} \leq P_{M_n}^{\check{\tau}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}}) \leq n^{-\lambda+\delta}\right) = 1. \quad (28)$$

#### 4.1 Proof of Proposition 15

Under the probability measure  $\mathbb{P}$ , we set, for every  $k \geq 1$ ,

$$p_k = p_k(\check{\tau}) := P_{M_k}^{\check{\tau}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k+1}}).$$

By the definition of  $M_k$ , there exists at least one plane tree grafted to  $\mathbf{u}_{M_k}$  that reaches generation 0. The root of this subtree is necessarily a child of  $\mathbf{u}_{M_k}$  distinct from  $\mathbf{u}_{M_{k-1}}$ . If such a subtree is unique, we let  $c_k = c_k(\check{\tau})$  be the probability that a simple random walk starting from its root reaches generation 0 before visiting  $\mathbf{u}_{M_k}$ . If there is more than one such grafted trees, we take  $c_k$  to be the sum of the corresponding probabilities. This definition is justified by the fact that  $c_k$  can be interpreted as the effective conductance between  $\mathbf{u}_{M_k}$  and generation 0 in the graph that consists only of the vertex  $\mathbf{u}_{M_k}$  and all the subtrees grafted to it.

We also set, for every  $k \geq 1$ ,

$$h_k = h_k(\check{\tau}) := P_{M_{k-1}}^{\check{\tau}}(S_0 < \Pi_{M_k}),$$

which is the probability that a simple random walk starting from  $\mathbf{u}_{M_{k-1}}$  reaches generation 0 before visiting  $\mathbf{u}_{M_k}$ . With the notation of Section 3.3, it is clear that

$$h_k = \mathcal{C}_{M_{k-1}}([\check{\tau}]^{M_k-1}).$$

We write  $\ell_k = 1/L_k = (M_k - M_{k-1})^{-1}$  for all  $k \geq 1$ . Then simple considerations show that

$$p_1 = \frac{\ell_1}{\ell_1 + c_1 + \ell_2},$$

and, for all  $k \geq 2$ ,

$$p_k = \frac{\ell_k}{\ell_k + c_k + \ell_{k+1}} \left( p_{k-1} + \frac{\ell_k}{\ell_k + c_{k-1} + h_{k-1}} p_k \right). \quad (29)$$

To establish the last formula, we consider the excursions of simple random walk outside of vertex  $\mathbf{u}_{M_k}$ , which are independent of the same law. Under this excursion law, the random walk makes its first jump with equal probability towards one of its neighbors, which are  $\mathbf{u}_{M_{k-1}}$ ,  $\mathbf{u}_{M_{k+1}}$  and the children of  $\mathbf{u}_{M_k}$  distinct from  $\mathbf{u}_{M_{k-1}}$ . The respective probabilities for an excursion to visit  $\mathbf{u}_{M_{k-1}}$ , to visit  $\mathbf{u}_{M_{k+1}}$ , and to reach generation 0 in one of the subtrees grafted at  $\mathbf{u}_{M_k}$ , are proportional respectively to  $1/L_k = \ell_k$ , to  $1/L_{k+1} = \ell_{k+1}$  and to  $c_k$ . So the probability for the random walk starting from  $\mathbf{u}_{M_k}$  to visit  $\mathbf{u}_{M_{k-1}}$  before hitting  $\mathbf{u}_{M_{k+1}}$  or reaching generation 0 is

$$\frac{\ell_k}{\ell_k + \ell_{k+1} + c_k}.$$

Next, conditionally on the latter event, the strong Markov property leads us to consider a simple random walk that starts from  $\mathbf{u}_{M_{k-1}}$ . With probability  $p_{k-1}$  it reaches generation 0 by hitting

the vertex  $\mathbf{u}_0$  before moving down to  $\mathbf{u}_{M_k}$ . However, we must also add the probability that this random walk goes back down to  $\mathbf{u}_{M_k}$  before reaching generation 0, which is equal to

$$\frac{\ell_k}{\ell_k + c_{k-1} + h_{k-1}},$$

multiplied by the probability  $p_k$  that once returning to  $\mathbf{u}_{M_k}$  the random walk will eventually hit generation 0 at  $\mathbf{u}_0$  before moving down to  $\mathbf{u}_{M_{k+1}}$ .

We derive from (29) that

$$p_{k-1} = p_k \left( \frac{\ell_k + c_k + \ell_{k+1}}{\ell_k} - \frac{\ell_k}{\ell_k + c_{k-1} + h_{k-1}} \right),$$

from which it follows that

$$p_1 = p_k \times \prod_{j=2}^k \left( 1 + \frac{c_j + \ell_{j+1}}{\ell_j} - \frac{\ell_j}{\ell_j + c_{j-1} + h_{j-1}} \right).$$

We define thus, for every  $j \geq 2$ ,

$$Q_j = Q_j(\check{\Gamma}) := \log \left( 1 + \frac{c_j + \ell_{j+1}}{\ell_j} - \frac{\ell_j}{\ell_j + c_{j-1} + h_{j-1}} \right).$$

**Lemma 16.** *We have*

$$\frac{1}{k} \sum_{j=2}^k Q_j \xrightarrow[k \rightarrow \infty]{L^2(\mathbb{P})} \frac{\lambda}{2}. \quad (30)$$

The proof of this key lemma is postponed to the next section. Let us first show how it implies Proposition 15 and thus Theorem 1. For any  $\delta > 0$ , consider the event

$$\left\{ (\lambda - \delta) \log n \leq \sum_{j=2}^{k_n} Q_j \leq (\lambda + \delta) \log n \right\}.$$

Using Lemma 13 and Lemma 16, we see that the last event holds with  $\mathbb{P}$ -probability tending to 1 as  $n \rightarrow \infty$ . As

$$p_{k_n} = p_1 \exp \left( - \sum_{j=2}^{k_n} Q_j \right),$$

we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( p_1 n^{-\lambda-\delta} \leq p_{k_n} \leq p_1 n^{-\lambda+\delta} \right) = 1.$$

Recalling the definition of  $p_k$ , we conclude that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( p_1 n^{-\lambda-\delta} \leq P_{M_{k_n}}^{\check{\Gamma}}(X_{S_0} = \mathbf{u}_0, S_0 < \Pi_{M_{k_n+1}}) \leq p_1 n^{-\lambda+\delta} \right) = 1.$$

Since  $\delta$  is arbitrary, the required convergence (28) readily follows from the last display. Therefore, it remains to prove Lemma 16.

## 4.2 Proof of Lemma 16

By the definition of  $Q_k$ , we can write for every  $k \geq 2$ ,

$$Q_k = \log \left( 1 + \frac{M_k c_k + \frac{M_k}{L_{k+1}}}{\frac{M_k}{L_k}} - \frac{\frac{M_k}{L_k}}{\frac{M_k}{L_k} + \frac{M_k}{M_{k-1}}(M_{k-1}c_{k-1} + M_{k-1}h_{k-1})} \right).$$

**Lemma 17.** *We have*

$$\left( \frac{L_{k+1}}{M_k}, \frac{L_k}{M_{k-1}}, M_k c_k, M_{k-1} c_{k-1}, M_{k-1} h_{k-1} \right) \xrightarrow[k \rightarrow \infty]{(d)} (\mathcal{R}, \mathcal{R}', \mathcal{C}, \mathcal{C}', \hat{\mathcal{C}}),$$

where in the limit

- $\mathcal{R}$  and  $\mathcal{R}'$  are two positive random variables with the same distribution given by

$$\mathbb{P}(\mathcal{R} > x) = (1+x)^{-2} \text{ for all } x \geq 0;$$

- $\mathcal{C}$  and  $\mathcal{C}'$  are distributed according to the law  $\gamma$ ;
- $\hat{\mathcal{C}}$  is distributed according to the law  $\hat{\gamma}$ .

Furthermore, we suppose that  $\mathcal{R}, \mathcal{R}', \mathcal{C}, \mathcal{C}', \hat{\mathcal{C}}$  are all defined under the probability measure  $\mathbb{P}$ , and they are independent.

*Proof.* We first observe that  $(M_{k-1}, L_k)_{k \geq 1}$  is a homogeneous Markov chain on  $\mathbb{Z}_+ \times \mathbb{N}$  whose initial distribution and transition probabilities are given as follows. Initially  $M_0 = 0$  and for every integer  $\ell \geq 1$ ,

$$\mathbb{P}(L_1 > \ell) = \prod_{j=1}^{\ell} \mathbb{P}(\epsilon_j = 0),$$

where  $\{\epsilon_j = 0\}$  means as previously that none of the Galton–Watson trees grafted at  $\mathbf{u}_j$  reaches generation 0. Then for every  $k \geq 1$ ,  $M_k = M_{k-1} + L_k$  and conditionally on  $\{M_k = m\}$ ,

$$\mathbb{P}(L_{k+1} > \ell \mid M_k = m) = \prod_{j=1}^{\ell} \mathbb{P}(\epsilon_{m+j} = 0) =: F(m, \ell), \quad \text{for every } \ell \geq 1. \quad (31)$$

Using (25), it is elementary to verify that for every  $x > 0$ ,

$$F(m, \lfloor xm \rfloor) \xrightarrow{m \rightarrow \infty} \frac{1}{(1+x)^2}.$$

For every  $k \geq 1$ , let  $\mathcal{F}_k = \sigma(L_1, L_2, \dots, L_k)$  be the  $\sigma$ -field generated by  $(L_i, 1 \leq i \leq k)$ , so that  $(\mathcal{F}_k)_{k \geq 1}$  is the natural filtration associated with the Markov chain  $(M_{k-1}, L_k)_{k \geq 1}$ . As  $M_k \geq k$ , it is clear that  $M_k \rightarrow \infty$  as  $k \rightarrow \infty$ . By dominated convergence and the last display, we get that for every  $x, y > 0$ ,

$$\begin{aligned} \mathbb{P}\left(\frac{L_k}{M_{k-1}} > x, \frac{L_{k+1}}{M_k} > y\right) &= \mathbb{E}\left[\mathbf{1}\left\{\frac{L_k}{M_{k-1}} > x\right\} \mathbb{P}\left(\frac{L_{k+1}}{M_k} > y \mid \mathcal{F}_k\right)\right] \\ &= \mathbb{E}\left[\mathbf{1}\left\{\frac{L_k}{M_{k-1}} > x\right\} \mathbb{P}\left(\frac{L_{k+1}}{M_k} > y \mid M_k\right)\right] \\ &\xrightarrow[k \rightarrow \infty]{} \frac{1}{(1+x)^2(1+y)^2}, \end{aligned}$$

which entails that

$$\left( \frac{L_k}{M_{k-1}}, \frac{L_{k+1}}{M_k} \right) \xrightarrow[k \rightarrow \infty]{(d)} (\mathcal{R}', \mathcal{R}). \quad (32)$$

Note that conditionally on  $M_k$  and on the number of subtrees grafted at  $\mathbf{u}_{M_k}$  that hit generation 0, these subtrees are Galton–Watson trees conditioned to have height greater than  $M_k - 1$ . Furthermore, the property  $\mathbb{E}[\hat{N}] = \sum k^2 \theta(k) < \infty$  and the estimate (21) entail that a.s. for all sufficiently large  $k$ , there is a unique subtree grafted at  $\mathbf{u}_{M_k}$  that reaches generation 0. Hence by Proposition 11, we obtain the convergence

$$(M_{k-1}c_{k-1}, M_k c_k) \xrightarrow[k \rightarrow \infty]{(d)} (\mathcal{C}', \mathcal{C}), \quad (33)$$

which holds jointly with (32), provided we let  $(\mathcal{C}', \mathcal{C})$  be independent of  $(\mathcal{R}', \mathcal{R})$ .

Let  $J \geq 2$  be a fixed integer. We can generalize the preceding arguments to show that the  $2J$ -tuple

$$\left( \frac{L_{k-1}}{M_{k-2}}, \frac{L_{k-2}}{M_{k-3}}, \dots, \frac{L_{k-J}}{M_{k-J-1}}, M_{k-2}c_{k-2}, M_{k-3}c_{k-3}, \dots, M_{k-J-1}c_{k-J-1} \right)$$

converges in distribution as  $k \rightarrow \infty$  to  $(\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_J, \mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_J)$ . These random variables appearing in the limit are all independent, and  $(\mathcal{R}_j)_{1 \leq j \leq J}$ , respectively  $(\mathcal{C}_j)_{1 \leq j \leq J}$ , have the same distribution as  $\mathcal{R}$ , resp. as  $\mathcal{C}$ . If we set  $V_j = \frac{\mathcal{R}_j}{1+\mathcal{R}_j}$  for all  $1 \leq j \leq J$ , then  $(V_j)_{1 \leq j \leq J}$  are i.i.d. with the same law of density  $2(1-x)$  on  $[0, 1]$ . Moreover, the previous convergence can be reformulated as

$$\left( \frac{L_{k-j}}{M_{k-j}}, M_{k-j-1}c_{k-j-1} \right)_{1 \leq j \leq J} \xrightarrow[k \rightarrow \infty]{(d)} (V_j, \mathcal{C}_j)_{1 \leq j \leq J}. \quad (34)$$

For all integers  $k > J$  and  $0 \leq j \leq J$ , we can define

$$h_k^{(j)} = h_k^{(j)}(\check{\tau}) := P_{M_{k-1}}^{\check{\tau}}(S_0 \wedge \Pi_{M_{k-j-1}} < \Pi_{M_k}),$$

which is the probability that a simple random walk starting from  $\mathbf{u}_{M_{k-1}}$  reaches generation 0 or the vertex  $\mathbf{u}_{M_{k-j-1}}$  before visiting  $\mathbf{u}_{M_k}$ . By definition it is clear that  $h_k^{(0)} = 1$ . From the interpretation of  $h_{k-1}^{(J)}$  as an electric conductance, we obtain the formula

$$h_{k-1}^{(J)} = \left( L_{k-1} + (c_{k-2} + h_{k-2}^{(J-1)})^{-1} \right)^{-1}$$

by the series law and parallel law. It follows that

$$M_{k-1}h_{k-1}^{(J)} = \left( \frac{L_{k-1}}{M_{k-1}} + \frac{1 - \frac{L_{k-1}}{M_{k-1}}}{M_{k-2}c_{k-2} + M_{k-2}h_{k-2}^{(J-1)}} \right)^{-1}.$$

The same calculation can be repeated for  $M_{k-2}h_{k-2}^{(J-1)}, M_{k-3}h_{k-3}^{(J-2)} \dots$  up until  $M_{k-J}h_{k-J}^{(1)}$ . By using (34) and the fact that  $h_{k-J-1}^{(0)} = 1$ , we see that the law of  $M_{k-1}h_{k-1}^{(J)}$  converges weakly to the  $J$ -th iterate  $\hat{\Phi}^J(\delta_\infty)$  as  $k \rightarrow \infty$ , where  $\hat{\Phi}$  is the mapping defined in (11). Moreover, according to assertion (1) in Proposition 7,  $\hat{\Phi}^J(\delta_\infty)$  converges weakly to  $\hat{\gamma} = \text{Law}(\hat{\mathcal{C}})$  as  $J \rightarrow \infty$ .

On the other hand,

$$0 \leq h_k^{(J)} - h_k = P_{M_{k-1}}^{\check{\tau}}(\Pi_{M_{k-J-1}} < \Pi_{M_k} < S_0).$$

Note that there is probability at least  $1 - \frac{M_{k-J-1}}{M_k}$  that, after hitting the vertex  $\mathbf{u}_{M_{k-J-1}}$ , the simple random walk on  $\check{\mathbf{T}}$  will reach generation 0 before hitting  $\mathbf{u}_{M_k}$ . Hence, by the strong Markov property of simple random walk, we have

$$h_k^{(J)} - h_k \leq P_{M_{k-1}}^{\check{\mathbf{T}}}(\Pi_{M_{k-J-1}} < S_0 \wedge \Pi_{M_k}) \times \frac{M_{k-J-1}}{M_k}.$$

By similar reasoning,

$$P_{M_{k-1}}^{\check{\mathbf{T}}}(\Pi_{M_{k-J-1}} < S_0 \wedge \Pi_{M_k}) \leq \frac{1}{M_k - M_{k-J} + 1},$$

and it follows that

$$M_k h_k^{(J)} - M_k h_k \leq \frac{M_{k-J} - 1}{M_k - M_{k-J} + 1}.$$

Thus, for any  $\eta > 0$ ,

$$\mathbb{P}(|M_{k-1} h_{k-1} - M_{k-1} h_{k-1}^{(J)}| \geq \eta) \leq \mathbb{P}\left(\frac{M_{k-J-1} - 1}{M_{k-1} - M_{k-J-1} + 1} \geq \eta\right).$$

But due to the previous discussions, it is clear that

$$\lim_{J \rightarrow \infty} \limsup_{k \rightarrow \infty} \mathbb{P}\left(\frac{M_{k-J-1} - 1}{M_{k-1} - M_{k-J-1} + 1} \geq \eta\right) = 0.$$

So we obtain, for any  $\eta > 0$ , that

$$\lim_{J \rightarrow \infty} \limsup_{k \rightarrow \infty} \mathbb{P}(|M_{k-1} h_{k-1} - M_{k-1} h_{k-1}^{(J)}| \geq \eta) = 0.$$

Finally, by applying [2, Theorem 3.2] we get that

$$M_{k-1} h_{k-1} \xrightarrow[k \rightarrow \infty]{(d)} \widehat{\mathcal{C}}. \quad (35)$$

Notice that  $h_{k-1}$  only depends on  $M_{k-1}$  and the finite tree strictly above the vertex  $\mathbf{u}_{M_{k-1}}$  in  $\check{\mathbf{T}}$ . So conditionally on  $M_{k-1}$ , the latter quantity  $h_{k-1}$  is independent of  $(L_k, L_{k+1}, c_{k-1}, c_k)$ . In consequence, the convergence (35) holds jointly with (32) and (33), if we take  $\widehat{\mathcal{C}}$  to be independent of  $(\mathcal{R}, \mathcal{R}', \mathcal{C}, \mathcal{C}')$ . The proof of Lemma 17 is therefore complete.  $\square$

**Lemma 18.** *It holds that*

$$\sup_{k \geq 1} \mathbb{E}[(M_k h_k)^2] < \infty.$$

*Proof.* From the interpretation of  $h_k$  as a conductance, we know by the series law and parallel law that

$$h_k = \left( L_k + (c_{k-1} + h_{k-1})^{-1} \right)^{-1}.$$

It follows that

$$M_k h_k = \frac{M_{k-1} c_{k-1} + M_{k-1} h_{k-1}}{\frac{L_k}{M_k} (M_{k-1} c_{k-1} + M_{k-1} h_{k-1}) + 1 - \frac{L_k}{M_k}}.$$

Again from the interpretation of  $c_{k-1}$  and  $h_{k-1}$  as conductances, it is elementary to see that  $M_{k-1} c_{k-1} \geq 1$  and  $M_{k-1} h_{k-1} \geq 1$ . Hence,

$$M_k h_k \leq \frac{M_{k-1} c_{k-1} + M_{k-1} h_{k-1}}{\frac{2L_k}{M_k} + 1 - \frac{L_k}{M_k}} \leq M_{k-1} c_{k-1} + \frac{M_{k-1} h_{k-1}}{1 + \frac{L_k}{M_k}}.$$

For any  $\eta > 0$ , one can take  $C(\eta) > 1 + \frac{1}{\eta}$  so that  $(a+b)^2 \leq C(\eta)a^2 + (1+\eta)b^2$  for every  $a, b > 0$ . Applying this inequality to the last display, we obtain

$$\mathbb{E}[(M_k h_k)^2] \leq C(\eta) \mathbb{E}[(M_{k-1} c_{k-1})^2] + (1+\eta) \mathbb{E}\left[\left(\frac{M_{k-1} h_{k-1}}{1 + \frac{L_k}{M_k}}\right)^2\right]. \quad (36)$$

Now notice that

$$\mathbb{E}\left[\left(\frac{M_{k-1} h_{k-1}}{1 + \frac{L_k}{M_k}}\right)^2\right] = \mathbb{E}\left[\sum_{\ell=1}^{\infty} \mathbb{P}(L_k = \ell | M_{k-1}) \frac{(M_{k-1} h_{k-1})^2}{(1 + \frac{\ell}{M_{k-1} + \ell})^2}\right].$$

According to (31), there exists a constant  $c > 0$  such that, a.s. for all integers  $k \geq 2$ , we have  $\mathbb{P}(L_k \geq M_{k-1} | M_{k-1}) \geq c$ . It follows that

$$\begin{aligned} & \sum_{\ell \geq 1} \mathbb{P}(L_k = \ell | M_{k-1}) \frac{(M_{k-1} h_{k-1})^2}{(1 + \frac{\ell}{M_{k-1} + \ell})^2} \\ & \leq \left(\frac{2}{3}\right)^2 \sum_{\ell \geq M_{k-1}} \mathbb{P}(L_k = \ell | M_{k-1}) (M_{k-1} h_{k-1})^2 + \sum_{\ell < M_{k-1}} \mathbb{P}(L_k = \ell | M_{k-1}) (M_{k-1} h_{k-1})^2 \\ & \leq (1 - \frac{5}{9}c) (M_{k-1} h_{k-1})^2, \end{aligned}$$

and thus

$$\mathbb{E}\left[\left(\frac{M_{k-1} h_{k-1}}{1 + \frac{L_k}{M_k}}\right)^2\right] \leq (1 - \frac{5}{9}c) \mathbb{E}[(M_{k-1} h_{k-1})^2].$$

Together with (36), the last display entails that

$$\mathbb{E}[(M_k h_k)^2] \leq C(\eta) \mathbb{E}[(M_{k-1} c_{k-1})^2] + (1+\eta)(1 - \frac{5}{9}c) \mathbb{E}[(M_{k-1} h_{k-1})^2].$$

Recall that by Lemma 12,  $\mathbb{E}[(M_{k-1} c_{k-1})^2]$  is uniformly bounded with respect to  $k$ . So by choosing  $\eta$  sufficiently small so that  $(1+\eta)(1 - \frac{5}{9}c) < 1$ , we see that there exist positive constants  $C < \infty$  and  $\rho < 1$ , both independent of  $k$ , such that for all  $k \geq 2$ ,

$$\mathbb{E}[(M_k h_k)^2] \leq C + \rho \mathbb{E}[(M_{k-1} h_{k-1})^2].$$

The sequence  $(\mathbb{E}[(M_k h_k)^2])_{k \geq 1}$  is therefore bounded.  $\square$

With the notation of Lemma 17, we set

$$Q_\infty := \log \left( 1 + \frac{\mathcal{C} + \frac{1}{\mathcal{R}}}{\frac{1+\mathcal{R}'}{\mathcal{R}'}} - \frac{1}{1 + \mathcal{R}'(\mathcal{C}' + \widehat{\mathcal{C}})} \right).$$

**Lemma 19.** (i) We have  $\lim_{k \rightarrow \infty} \mathbb{E}[Q_k] = \mathbb{E}[Q_\infty]$ .

(ii) We have the equality  $\mathbb{E}[Q_\infty] = \lambda/2$ .

(iii) It holds that

$$\sup_{i,j \geq 1} \mathbb{E}[|Q_i Q_j|] < \infty \quad \text{and} \quad \sup_{i,j \geq 1} \mathbb{E}[(Q_i Q_j)^2] < \infty.$$

*Proof.* (i) On the one hand, since  $\frac{M_k}{L_k} > 1$ ,

$$Q_k \leq \log \left( 1 + \frac{M_k c_k + \frac{M_k}{L_{k+1}}}{\frac{M_k}{L_k}} \right) \leq \log \left( 1 + M_k c_k + \frac{M_k}{L_{k+1}} \right).$$

On the other hand,

$$\begin{aligned} Q_k &\geq \log \left( 1 - \frac{\frac{M_k}{L_k}}{\frac{M_k}{L_k} + \frac{M_k}{M_{k-1}}(M_{k-1}c_{k-1} + M_{k-1}h_{k-1})} \right) \\ &= -\log \left( 1 + \frac{\frac{M_k}{L_k}}{\frac{M_k}{M_{k-1}}(M_{k-1}c_{k-1} + M_{k-1}h_{k-1})} \right) = -\log \left( 1 + \frac{\frac{M_{k-1}}{L_k}}{M_{k-1}c_{k-1} + M_{k-1}h_{k-1}} \right). \end{aligned}$$

Using the facts that  $M_{k-1}c_{k-1} \geq 1$  and  $M_{k-1}h_{k-1} \geq 1$ , we arrive at

$$|Q_k| \leq \max \left\{ \log \left( 1 + M_k c_k + \frac{M_k}{L_{k+1}} \right), \log \left( 1 + \frac{M_{k-1}}{2L_k} \right) \right\}.$$

One can find  $A > 0$  such that  $\log(1+x) \leq A + x^{1/2}$  for every  $x > 0$ . It follows that

$$|Q_k| \leq A + \left( M_k c_k + \frac{M_k}{L_{k+1}} \right)^{\frac{1}{2}} + \left( \frac{M_{k-1}}{L_k} \right)^{\frac{1}{2}} \leq A + (M_k c_k)^{\frac{1}{2}} + \left( \frac{M_k}{L_{k+1}} \right)^{\frac{1}{2}} + \left( \frac{M_{k-1}}{L_k} \right)^{\frac{1}{2}}. \quad (37)$$

Recall the convergence in distribution  $M_k c_k \xrightarrow{(d)} \mathcal{C}$  shown in Lemma 17. By Lemma 12, it follows that

$$\mathbb{E}[M_k c_k] \xrightarrow[k \rightarrow \infty]{} \mathbb{E}[\mathcal{C}]. \quad (38)$$

In particular,  $\sup_k \mathbb{E}[M_k c_k] < \infty$ . Meanwhile, using (25) and (31), it is not difficult to verify that there exists a positive constant  $K$  such that for every  $x > 0$ ,

$$\sup_{k \geq 1} \mathbb{P} \left( \frac{M_k}{L_{k+1}} > x \right) \leq \frac{K}{1+x}.$$

So using the formula

$$\mathbb{E} \left[ \left( \frac{M_k}{L_{k+1}} \right)^\alpha \right] = \int_0^\infty \alpha x^{\alpha-1} \mathbb{P} \left( \frac{M_k}{L_{k+1}} > x \right) dx,$$

we get the existence of a constant  $\alpha \in (\frac{1}{2}, 1)$  such that

$$\sup_{k \geq 1} \mathbb{E} \left[ \left( \frac{M_k}{L_{k+1}} \right)^\alpha \right] < \infty. \quad (39)$$

Hence, it follows from (37) that  $(Q_k)_{k \geq 2}$  is bounded in  $L^p$  with some  $p > 1$ . The sequence  $(Q_k)_{k \geq 2}$  is thus uniformly integrable. However, according to Lemma 17,  $Q_k$  converges in distribution to  $Q_\infty$ . Therefore,  $Q_k \rightarrow Q_\infty$  in  $L^1$  and we have

$$\lim_{k \rightarrow \infty} \mathbb{E}[Q_k] = \mathbb{E}[Q_\infty].$$



(ii) Recall that  $V = \frac{\mathcal{R}}{1+\mathcal{R}}$  and  $V' = \frac{\mathcal{R}'}{1+\mathcal{R}'}$  are independent with the same law of density  $2(1-x)$  on  $[0, 1]$ . Noting that

$$\begin{aligned} \log \left( 1 + \frac{\mathcal{C} + \frac{1}{\mathcal{R}}}{\frac{1+\mathcal{R}'}{\mathcal{R}'}} - \frac{1}{1 + \mathcal{R}'(\mathcal{C}' + \widehat{\mathcal{C}})} \right) &= \log \left( \left( \mathcal{C} + \frac{1}{\mathcal{R}} \right) V' + \frac{V'(\mathcal{C}' + \widehat{\mathcal{C}})}{1 - V' + V'(\mathcal{C}' + \widehat{\mathcal{C}})} \right) \\ &= \log V' + \log \left( \mathcal{C} + \frac{1}{\mathcal{R}} + \left( V' + \frac{1 - V'}{\mathcal{C}' + \widehat{\mathcal{C}}} \right)^{-1} \right), \end{aligned}$$

we can use the distributional identity (6) to obtain

$$\mathbb{E}[Q_\infty] = \mathbb{E}[\log V'] + \mathbb{E} \left[ \log \left( \mathcal{C} + \frac{1}{\mathcal{R}} + \widehat{\mathcal{C}} \right) \right].$$

Since  $\mathcal{R} = \frac{V}{1-V}$ , it follows that

$$\log \left( \mathcal{C} + \frac{1}{\mathcal{R}} + \widehat{\mathcal{C}} \right) = \log(1 - V + V(\mathcal{C} + \widehat{\mathcal{C}})) - \log V,$$

which yields  $\mathbb{E}[Q_\infty] = \mathbb{E}[\log(1 - V + V(\mathcal{C} + \widehat{\mathcal{C}}))]$ . To complete the proof of assertion (ii), we apply (6) again to see that

$$\mathbb{E}[\log(1 - V + V(\mathcal{C} + \widehat{\mathcal{C}}))] = \mathbb{E}[\log(\widehat{\mathcal{C}} + \mathcal{C})] - \mathbb{E}[\log(\widehat{\mathcal{C}})] = \mathbb{E} \left[ \log \left( \frac{\widehat{\mathcal{C}} + \mathcal{C}}{\widehat{\mathcal{C}}} \right) \right],$$

which is equal to  $\lambda/2$  according to (20).

(iii) There exists a constant  $\tilde{A} > 0$  such that  $\log(1+x) \leq \tilde{A} + x^{1/4}$  for every  $x > 0$ . It follows then from the same arguments as in the proof of assertion (i) that

$$\begin{aligned} |Q_i Q_j| &\leq \left( \log \left( 1 + M_i c_i + \frac{M_i}{L_{i+1}} \right) + \log \left( 1 + \frac{M_{i-1}}{2L_i} \right) \right) \left( \log \left( 1 + M_j c_j + \frac{M_j}{L_{j+1}} \right) + \log \left( 1 + \frac{M_{j-1}}{2L_j} \right) \right) \\ &\leq \left( 2\tilde{A} + \left( M_i c_i + \frac{M_i}{L_{i+1}} \right)^{\frac{1}{4}} + \left( \frac{M_{i-1}}{L_i} \right)^{\frac{1}{4}} \right) \left( 2\tilde{A} + \left( M_j c_j + \frac{M_j}{L_{j+1}} \right)^{\frac{1}{4}} + \left( \frac{M_{j-1}}{L_j} \right)^{\frac{1}{4}} \right). \end{aligned}$$

In order to prove that  $\sup_{i,j} \mathbb{E}[|Q_i Q_j|] < \infty$ , it is enough to develop the product in the last line of the preceding display, and to show that the expectation of each term is uniformly bounded with respect to  $i$  and  $j$ . In fact, by the Cauchy-Schwarz inequality,

$$\mathbb{E} \left[ \left( \frac{M_{i-1}}{L_i} \right)^{\frac{1}{4}} \left( \frac{M_{j-1}}{L_j} \right)^{\frac{1}{4}} \right] \leq \mathbb{E} \left[ \left( \frac{M_{i-1}}{L_i} \right)^{\frac{1}{2}} \right]^{\frac{1}{2}} \mathbb{E} \left[ \left( \frac{M_{j-1}}{L_j} \right)^{\frac{1}{2}} \right]^{\frac{1}{2}},$$

in which the right-hand side is uniformly bounded according to (39). Moreover, as

$$\mathbb{E} \left[ \left( M_i c_i + \frac{M_i}{L_{i+1}} \right)^{\frac{1}{4}} \left( M_j c_j + \frac{M_j}{L_{j+1}} \right)^{\frac{1}{4}} \right] \leq \mathbb{E} \left[ \left( (M_i c_i)^{\frac{1}{4}} + \left( \frac{M_i}{L_{i+1}} \right)^{\frac{1}{4}} \right) \left( (M_j c_j)^{\frac{1}{4}} + \left( \frac{M_j}{L_{j+1}} \right)^{\frac{1}{4}} \right) \right],$$

we can develop the right-hand side and similarly use (38) and (39) to show that it is uniformly bounded. All the other terms can be treated in an analogous way. By similar arguments, one can also prove that  $\sup_{i,j} \mathbb{E}[(Q_i Q_j)^2] < \infty$ . This finishes the proof of assertion (iii).  $\square$

Using assertions (i) and (ii) of Lemma 19, we have thus

$$\lim_{k \rightarrow \infty} \mathbb{E} \left[ \frac{1}{k} \sum_{j=2}^k Q_j \right] = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=2}^k \mathbb{E}[Q_j] = \mathbb{E}[Q_\infty] = \frac{\lambda}{2}. \quad (40)$$

**Lemma 20.** *We have*

$$\limsup_{k \rightarrow \infty} \mathbb{E} \left[ \left( \frac{1}{k} \sum_{j=2}^k Q_j \right)^2 \right] \leq (\mathbb{E}[Q_\infty])^2 = \frac{\lambda^2}{4}. \quad (41)$$

*Proof.* We first note that for every  $\delta \in (0, 1/2)$ ,

$$\left| \mathbb{E} \left[ \sum_{2 \leq i, j \leq k} Q_i Q_j \right] - \mathbb{E} \left[ \sum_{\substack{\delta k \leq i, j \leq k \\ |i-j| > \delta k}} Q_i Q_j \right] \right| \leq 4\delta k^2 \times \sup_{i,j} \mathbb{E}[|Q_i Q_j|].$$

In view of Lemma 19 (iii), the estimate (41) will be proved if we can show for arbitrarily small  $\delta \in (0, 1/2)$  that

$$\limsup_{k \rightarrow \infty} \frac{1}{k^2} \mathbb{E} \left[ \sum_{\substack{\delta k \leq i, j \leq k \\ |i-j| > \delta k}} Q_i Q_j \right] \leq (\mathbb{E}[Q_\infty])^2. \quad (42)$$

We thus fix  $\delta \in (0, 1/2)$  in the following arguments. By symmetry, we can further restrict our attention to the indices  $i$  and  $j$  such that  $\delta k \leq i, j \leq k$  and  $j - i > \delta k$ .

Notice that by the Cauchy–Schwarz inequality,

$$\begin{aligned} \left| \frac{1}{k^2} \mathbb{E} \left[ \sum_{\substack{\delta k \leq i, j \leq k \\ j-i > \delta k}} Q_i Q_j \mathbf{1}_{\{M_{i+1} \geq \varepsilon M_{j-1}\}} \right] \right| &\leq \frac{1}{k^2} \sum_{\substack{\delta k \leq i, j \leq k \\ j-i > \delta k}} \mathbb{P}(M_{i+1} \geq \varepsilon M_{j-1})^{\frac{1}{2}} \mathbb{E}[(Q_i Q_j)^2]^{\frac{1}{2}} \\ &\leq \sup_{\substack{\delta k \leq i, j \leq k \\ j-i > \delta k}} \mathbb{P}(M_{i+1} \geq \varepsilon M_{j-1})^{\frac{1}{2}} \times \sup_{i,j} \mathbb{E}[(Q_i Q_j)^2]^{\frac{1}{2}}. \end{aligned} \quad (43)$$

However, observe that Lemma 13 can be reformulated as

$$\frac{\log M_k}{k} \xrightarrow[k \rightarrow \infty]{\mathbb{P}\text{-a.s.}} \frac{1}{2},$$

and it follows that for all  $\varepsilon \in (0, 1)$ ,

$$\lim_{k \rightarrow \infty} \mathbb{P}(\text{there exist } i, j \in [\delta k, k] \text{ with } j - i > \delta k \text{ such that } M_{i+1} \geq \varepsilon M_{j-1}) = 0. \quad (44)$$

Together with Lemma 19 (iii), the latter display implies that the right-hand side of (43) converges to 0 as  $k \rightarrow \infty$ . The proof of (42) is thus reduced to showing that for fixed  $\delta$ ,

$$\limsup_{\varepsilon \rightarrow 0} \left( \limsup_{k \rightarrow \infty} \frac{2}{k^2} \mathbb{E} \left[ \sum_{\substack{\delta k \leq i, j \leq k \\ j-i > \delta k}} Q_i Q_j \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}} \right] \right) \leq (\mathbb{E}[Q_\infty])^2. \quad (45)$$

To this end, we take  $\varepsilon \in (0, 1/2)$  and define, for every  $k \geq 2$ ,

$$h_k^\varepsilon = h_k^\varepsilon(\check{\mathbf{T}}) := \mathcal{C}_{M_k - 1 - \lfloor \varepsilon M_k \rfloor}([\check{\mathbf{T}}]^{*(M_k - 1)})$$

where  $[\check{\mathbf{T}}]^{*(M_k - 1)}$  stands for the reduced tree associated with  $[\check{\mathbf{T}}]^{M_k - 1}$  up to height  $M_k - 1$ . In other words,  $h_k^\varepsilon$  is the probability that a simple random walk on  $\check{\mathbf{T}}$  starting from  $\mathbf{u}_{M_k - 1}$  hits a point of generation  $-\lfloor \varepsilon M_k \rfloor$  that has a descendant at generation 0 before hitting  $\mathbf{u}_{M_k}$ .

Comparing with the definition of  $h_k$ , it is clear that  $h_k^\varepsilon \geq h_k$ . On the other hand, by similar arguments as in the proof of Proposition 11, we obtain

$$h_k^\varepsilon - h_k \leq \frac{\lfloor \varepsilon M_k \rfloor}{M_k} h_k^\varepsilon,$$

which entails that

$$M_k h_k^\varepsilon - M_k h_k \leq \lfloor \varepsilon M_k \rfloor h_k^\varepsilon \leq 2\varepsilon M_k h_k. \quad (46)$$

We set

$$Q_k^\varepsilon = Q_k^\varepsilon(\check{\mathbf{T}}) := \log \left( 1 + \frac{M_k c_k + \frac{M_k}{L_{k+1}}}{\frac{M_k}{L_k}} - \frac{\frac{M_k}{L_k}}{\frac{M_k}{L_k} + \frac{M_k}{M_{k-1}}(M_{k-1} c_{k-1} + M_{k-1} h_{k-1}^\varepsilon)} \right) \geq Q_k.$$

Using the elementary inequality  $0 \leq \log x - \log y \leq \frac{x-y}{y}$  for  $x \geq y > 0$ , we see that

$$\begin{aligned} Q_k^\varepsilon - Q_k &\leq \frac{\frac{M_k}{L_k}}{M_k c_k + \frac{M_k}{L_{k+1}}} \left( \frac{\frac{M_k}{L_k}}{\frac{M_k}{L_k} + \frac{M_k}{M_{k-1}}(M_{k-1} c_{k-1} + M_{k-1} h_{k-1})} - \frac{\frac{M_k}{L_k}}{\frac{M_k}{L_k} + \frac{M_k}{M_{k-1}}(M_{k-1} c_{k-1} + M_{k-1} h_{k-1}^\varepsilon)} \right) \\ &\leq \frac{\left(\frac{M_k}{L_k}\right)^2}{M_k c_k + \frac{M_k}{L_{k+1}}} \cdot \frac{\frac{M_k}{M_{k-1}}(M_{k-1} h_{k-1}^\varepsilon - M_{k-1} h_{k-1})}{\left(\frac{M_k}{L_k} + \frac{M_k}{M_{k-1}}(M_{k-1} c_{k-1} + M_{k-1} h_{k-1})\right)^2}. \end{aligned}$$

Taking account of the easy facts that  $M_k c_k \geq 1$  and  $M_k = M_{k-1} + L_k$ , we obtain

$$Q_k^\varepsilon - Q_k \leq \frac{\left(\frac{M_k}{L_k}\right)^2 \frac{M_k}{M_{k-1}}(M_{k-1} h_{k-1}^\varepsilon - M_{k-1} h_{k-1})}{\left(\frac{M_k}{L_k} + \frac{M_k}{M_{k-1}}\right)^2} \leq M_{k-1} h_{k-1}^\varepsilon - M_{k-1} h_{k-1},$$

which, together with (46), implies that

$$Q_k^\varepsilon - Q_k \leq 2\varepsilon M_{k-1} h_{k-1}.$$

This allows us to approximate  $\mathbb{E}[Q_i Q_j \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}}]$  by  $\mathbb{E}[Q_i Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}}]$ , because

$$\begin{aligned} \left| \mathbb{E}[Q_i Q_j \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}}] - \mathbb{E}[Q_i Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}}] \right| &\leq \mathbb{E}[|Q_i(Q_j^\varepsilon - Q_j)|] \\ &\leq \mathbb{E}[(Q_j^\varepsilon - Q_j)^2]^{\frac{1}{2}} \times \mathbb{E}[(Q_i)^2]^{\frac{1}{2}} \\ &\leq 2\varepsilon \mathbb{E}[(M_{j-1} h_{j-1})^2]^{\frac{1}{2}} \times \mathbb{E}[(Q_i)^2]^{\frac{1}{2}}, \end{aligned}$$

and the right-hand side converges to 0 uniformly with respect to  $i, j$  and  $k$  when  $\varepsilon \rightarrow 0$ , according to Lemma 18 and assertion (iii) of Lemma 19.

Let us return to the indices  $i, j$  such that  $\delta k \leq i, j \leq k$  and  $j - i > \delta k$ , and let  $\tilde{\mathcal{F}}_i$  be the  $\sigma$ -field generated by the variable  $M_{i+1}$  and the finite part of  $\check{\mathbf{T}}$  above the vertex  $\mathbf{u}_{M_{i+1}}$ . Informally, one can think of it as the  $\sigma$ -field generated by  $[\check{\mathbf{T}}]^{M_{i+1}}$ . Then  $Q_i$  is  $\tilde{\mathcal{F}}_i$ -measurable and

$$\mathbb{E}[Q_i Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}}] = \mathbb{E}\left[Q_i \mathbb{E}\left[Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}} \mid \tilde{\mathcal{F}}_i\right]\right]. \quad (47)$$

At this point, we observe that

$$\mathbb{E}\left[Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}} \mid \tilde{\mathcal{F}}_i\right] = \mathbb{E}\left[Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}} \mid M_{i+1}\right]. \quad (48)$$

On the other hand, one can generalize the proof of Lemma 17 to show that

$$\left(\frac{L_{j+1}}{M_j}, \frac{L_j}{M_{j-1}}, M_j c_j, M_{j-1} c_{j-1}, M_{j-1} h_{j-1}^\varepsilon\right) \xrightarrow{j \rightarrow \infty}^{(d)} (\mathcal{R}, \mathcal{R}', \mathcal{C}, \mathcal{C}', \widehat{\mathcal{C}}_\varepsilon), \quad (49)$$

where in the limit, the first four random variables  $\mathcal{R}, \mathcal{R}', \mathcal{C}, \mathcal{C}'$  are the same as in Lemma 17, whereas  $\widehat{\mathcal{C}}_\varepsilon$  is distributed as  $\mathcal{C}(\widehat{\Delta}_\varepsilon)$ . It is assumed in addition that  $\mathcal{R}, \mathcal{R}', \mathcal{C}, \mathcal{C}', \widehat{\mathcal{C}}_\varepsilon$  are independent under  $\mathbb{P}$ . Furthermore, we can verify that the convergence (49) is still valid if, instead of the distribution of

$$\left(\frac{L_{j+1}}{M_j}, \frac{L_j}{M_{j-1}}, M_j c_j, M_{j-1} c_{j-1}, M_{j-1} h_{j-1}^\varepsilon\right),$$

we consider the conditional distribution of the same random 5-tuple given  $M_{i+1}$ , and let  $i$  and  $j$  tend to infinity satisfying that  $j - i > \delta j$ .

We define thus

$$Q_\infty^\varepsilon := \log \left( 1 + \frac{\mathcal{C} + \frac{1}{\mathcal{R}}}{\frac{1 + \mathcal{R}'}{\mathcal{R}'}} - \frac{1}{1 + \mathcal{R}'(\mathcal{C}' + \widehat{\mathcal{C}}_\varepsilon)} \right),$$

which has the limiting distribution of the sequence  $(Q_j^\varepsilon)$  conditioned on  $M_{i+1}$ . By similar arguments used in the proof of assertion (i) of Lemma 19, we know that there exists some  $p > 1$  such that

$$\sup_{\ell: \mathbb{P}(M_{i+1}=\ell)>0} \mathbb{E}[(Q_j^\varepsilon)^p | M_{i+1} = \ell]$$

is uniformly bounded for all  $i, j$  satisfying that  $j - i > \delta j$ , and that

$$\lim_{\substack{i, j \rightarrow \infty \\ j-i > \delta j}} \left( \sup_{\ell: \mathbb{P}(M_{i+1}=\ell)>0} \left| \mathbb{E}[Q_j^\varepsilon | M_{i+1} = \ell] - \mathbb{E}[Q_\infty^\varepsilon] \right| \right) = 0.$$

In view of (44) and (48), it follows that a.s.

$$\lim_{\substack{i, j \rightarrow \infty \\ j-i > \delta j}} \left| \mathbb{E}[Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}} | \mathcal{F}_i] - \mathbb{E}[Q_\infty^\varepsilon] \right| = 0.$$

Hence, we get from (47) and Lemma 19 (i) that

$$\lim_{k \rightarrow \infty} \left( \sup_{\substack{i, j \in [\delta k, k] \\ j-i > \delta k}} \mathbb{E}[Q_i Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}}] \right) \leq \mathbb{E}[Q_\infty] \mathbb{E}[Q_\infty^\varepsilon].$$

Finally, it remains to estimate the difference between  $\mathbb{E}[Q_\infty^\varepsilon]$  and  $\mathbb{E}[Q_\infty]$ . To do this, we use a coupling argument by defining both  $\widehat{\mathcal{C}}_\varepsilon$  and  $\widehat{\mathcal{C}}$  from a common reduced tree  $\widehat{\Delta}$ , independent of  $(\mathcal{R}, \mathcal{R}', \mathcal{C}, \mathcal{C}')$ , so that  $\widehat{\mathcal{C}} = \mathcal{C}(\widehat{\Delta})$  and  $\widehat{\mathcal{C}}_\varepsilon = \mathcal{C}(\widehat{\Delta}_\varepsilon)$ . Since  $\mathcal{C} \geq 1$  and  $\widehat{\mathcal{C}}_\varepsilon \geq \widehat{\mathcal{C}} \geq 1$ , one can proceed in the same way as we did for bounding  $Q_k^\varepsilon - Q_k$ , and arrive at

$$0 \leq Q_\infty^\varepsilon - Q_\infty \leq \frac{1 + \frac{1}{\mathcal{R}'}}{\mathcal{C} + \frac{1}{\mathcal{R}}} \frac{\mathcal{R}'(\widehat{\mathcal{C}}_\varepsilon - \widehat{\mathcal{C}})}{(1 + \mathcal{R}'(\mathcal{C}' + \widehat{\mathcal{C}}))^2} \leq \left(1 + \frac{1}{\mathcal{R}'}\right) \frac{\mathcal{R}'(\widehat{\mathcal{C}}_\varepsilon - \widehat{\mathcal{C}})}{(1 + \mathcal{R}')^2} \leq \widehat{\mathcal{C}}_\varepsilon - \widehat{\mathcal{C}}.$$

Taking account of (22), the last display gives  $|\mathbb{E}[Q_\infty^\varepsilon] - \mathbb{E}[Q_\infty]| \leq 2\varepsilon \mathbb{E}[\widehat{\mathcal{C}}]$ , and the right-hand side converges to 0 as  $\varepsilon \rightarrow 0$ .

According to the previous discussions, we conclude that

$$\begin{aligned}
& \limsup_{\varepsilon \rightarrow 0} \left( \limsup_{k \rightarrow \infty} \frac{2}{k^2} \mathbb{E} \left[ \sum_{\substack{\delta k \leq i, j \leq k \\ j-i > \delta k}} Q_i Q_j \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}} \right] \right) \\
&= \limsup_{\varepsilon \rightarrow 0} \left( \limsup_{k \rightarrow \infty} \frac{2}{k^2} \mathbb{E} \left[ \sum_{\substack{\delta k \leq i, j \leq k \\ j-i > \delta k}} Q_i Q_j^\varepsilon \mathbf{1}_{\{M_{i+1} < \varepsilon M_{j-1}\}} \right] \right) \\
&\leq \limsup_{\varepsilon \rightarrow 0} \mathbb{E}[Q_\infty] \mathbb{E}[Q_\infty^\varepsilon] = (\mathbb{E}[Q_\infty])^2,
\end{aligned}$$

which finishes the proof of (45). The proof of Lemma 20 is therefore completed.  $\square$

*Proof of Lemma 16.* By combining (40) and (41), we have

$$\limsup_{k \rightarrow \infty} \mathbb{E} \left[ \left( \frac{1}{k} \sum_{j=2}^k Q_j - \frac{\lambda}{2} \right)^2 \right] \leq \left( \limsup_{k \rightarrow \infty} \mathbb{E} \left[ \left( \frac{1}{k} \sum_{j=2}^k Q_j \right)^2 \right] \right) - \lambda \lim_{k \rightarrow \infty} \mathbb{E} \left[ \frac{1}{k} \sum_{j=2}^k Q_j \right] + \frac{\lambda^2}{4} \leq 0,$$

which gives the desired result.

## References

- [1] K.B. ATHREYA, P. E. NEY, *Branching Processes*. Springer, New York–Heidelberg, 1972.
- [2] P. BILLINGSLEY, *Convergence of probability measures*, 2nd edition, John Wiley & Sons Inc., New York, 1999.
- [3] N. CURIEN, J.-F. LE GALL, *The harmonic measure of balls in random trees*, to appear in Ann. Probab., available at [arxiv:1304.7190](https://arxiv.org/abs/1304.7190).
- [4] T. DUQUESNE, J.-F. LE GALL, *The Hausdorff measure of stable trees*, Alea, 1 (2006), 393–415.
- [5] H. KESTEN, *Subdiffusive behavior of random walk on a random cluster*, Ann. Inst. H. Poincaré Probab. Statist., 22 (1986), 425–487.
- [6] S. LIN, *The harmonic measure of balls in critical Galton–Watson trees with infinite variance offspring distribution*. Electron. J. Probab. 19 (2014), no. 98, 1–35.
- [7] Q. LIU, A. ROUAULT, *On two measures defined on the boundary of a branching tree*, in Classical and Modern Branching Processes, The IMA Volumes in Mathematics and its Applications Volume 84, Springer New York, 1997, 187–201.
- [8] R. LYONS, R. PEMANTLE, AND Y. PERES, *Ergodic theory on Galton–Watson trees: speed of random walk and dimension of harmonic measure*, Ergodic Theory Dynam. Systems, 15 (1995), 593–619.
- [9] R. LYONS, R. PEMANTLE, AND Y. PERES, *Conceptual proofs of  $l \log l$  criteria for mean behavior of branching processes*, Ann. Probab., 23 (1995), 1125–1138.

- [10] R. LYONS, Y. PERES, *Probability on Trees and Networks*, Cambridge University Press, to appear, available at <http://pages.iu.edu/~rdlyons/>.
- [11] Y. PERES, O. ZEITOUNI, *A central limit theorem for biased random walks on Galton-Watson trees*, Probab. Theory Related Fields, 140 (2008), no. 3-4, 595–629.